DOCUMENT RESUME

ED 325 821 CS 010 308

AUTHOR	Harste, Jerome C., Ed.; Carey, Robert F., Ed.
TITLE	New Perspectives on Comprehension. Monographs in Teaching and Learning Number 3.
INSTITUTION	Indiana Univ., Bloomington. School of Education.
REPORT NO	ISSN-0193-4740
PUB DATE	Oct 79
NOTE	175p.; For monographs 1 and 4-5, see CS 010 307-310; for monograph 2, see ED 189 590; for monograph 6, see ED 323 564.
AVAILABLE FROM	ERIC Clearinghouse on Reading and Communication Skills, Indiana University, 2805 E. 10th St., Suite 150, Bloomington, IN 47408-2698 ($6.95 plus postage and handling).
PUB TYPE	Collected Works - General (020) -- Viewpoints (120)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Beginning Reading; Elementary Secondary Education; Evaluation Methods; Higher Education; Interdisciplinary Approach; *Reading Comprehension; *Reading Instruction; *Reading Research; Research Methodology
IDENTIFIERS	Meaning Conditions

ABSTRACT

Focusing on providing an interdisciplinary framework, the papers in this monograph attempt to synthesize knowledge about comprehension from the points of view of cognitive psychology, textlinguistics, artificial intelligence, and sociolinguistics, as well as from the more conventional perspective of language education. Papers include "Comprehension as Setting" (Jerome Harste and Robert Carey); "Toward Understanding Comprehension: An Examination of Systems" (Robert J. Tierney and others); "Reading is Recognition--When Reading is Not Reasoning" (J. Jaap Tuinman); "Overt Indications of Reading Comprehension: Product and Process Considerations" (William D. Page and Richard T. Vacca); "Cloze Restoration, Recall and Discourse Processing in Middle Grade Readers" (Sandra M. Gould); "Measures of Comprehension: Independence and Implications" (Pat Rigg and Elizabeth Taylor); "Retellings as Measures of Comprehension: A Perspective" (Sharon L. Smith); "Comprehension and Predictability: An Analysis of Beginning Reading Materials" (Lynn K. Rhodes); "Some Basic Notions about Reading Comprehension: Implications for Teachers" (Robert J. Tierney and Rand J. Spiro); "Bridging the Gaps in Reading: Respect and Communication" (Kenneth S. Goodman); and "Methodological Concerns in Comprehension Research" (P. David Pearson). (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

monograph
in language and reading studies

new perspectives on comprehension

Monographs in Teaching and Learning / School of Education / Indiana University
Number 3, October, 1979

ED325821

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Leo Fay

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS010308

monograph in language and reading studies

US ISSN 0193-4759 / Number 3 / October, 1979

new perspectives on comprehension

Jerome C. Harste
Robert F. Carey

Monographs in Teaching and Learning / School of Education / Indiana University
US ISSN 0193-4740

CS010368

Indiana University's Language Education Departments initiated the Monographs in Language and Reading Studies to serve a number of purposes, including the publication of papers based on conference proceedings, research and development reports, and papers of special interest to teachers and other professionals working in the language education areas. In a reorganization of the School of Education, the departments of Applied Linguistics, English Education, Foreign Language Education, and Reading Education were confederated into a new unit—the Language Education Department. Each unit maintains its identity and at the same time participates in mutual programs, research and development, and service activities. The Monographs are an outcome of this confederation.

Richard P. Gousha, Dean

Albert H. Fink, Editor; **Dorothy J. Taylor**, Associate Editor; **Carrie L. Anglin**, Assistant Editor; **Terri R. Burks**, Circulation

Monograph in Language and Reading Studies

Series Editor

Leo C. Fay, Reading Education

Editorial Board

Harry L. Gradman, Applied Linguistics; **Robert C. Lafayette**, Foreign Language Education; **Josephine Spear**, English Education

Monographs in Teaching and Learning comprise series of monographs in a variety of disciplines. Each series is published on an irregular schedule by Indiana University School of Education. Articles are devoted to aspects of major current concern to educators in the respective fields and deal with theory, research, or practical application pertaining to all levels of practice.

Contents

Introduction

Comprehension as Setting
Harste & Carey / 4

Toward Understanding Comprehension: An Examination of Systems
Tierney, Vaughan, & Bridge / 23

Reading Is Recognition—When Reading is not Reasoning
Tuinman / 38

Overt Indications of Reading Comprehension: Product and Process Consideration
Page & Vacca / 49

Cloze Restoration, Recall and Discourse Processing in Middle Grade Readers
Gould / 64

Measures of Comprehension: Independence and Implications
Rigg & Taylor / 77

Retellings as Measures of Comprehension: A Perspective
Smith / 84

Comprehension and Predictability: An Analysis of Beginning Reading Materials
Rhodes / 100

Some Basic Notions about Reading Comprehension: Implications for Teachers
Tierney & Spiro / 132

Bridging the Gaps in Reading: Respect and Communication
Goodman / 138

Methodological Concerns in Comprehension Research
Pearson / 153

Introduction

The information explosion in reading education has begun. The past decade witnessed an unprecedented quantity of theory and research into the basic process of reading and into the pedagogical strategies that most efficiently support reading.

During the same time span it has become clear to nearly everyone that comprehension—the search for and construction of meaning—is the heart of the matter. With this discovery has come an admirable, though at times misdirected, attempt to exert the force of a multidisciplinary perspective on comprehension. As Goodman suggests later in this Monograph, what is needed is an interdisciplinary, rather than a multidisciplinary, approach. *New Perspectives on Comprehension* is an attempt to provide a framework for such an approach. The volume attempts to synthesize current knowledge about comprehension from the points of view of cognitive psychology, textlinguistics, artificial intelligence, and sociolinguistics, as well as from the more conventional perspective of language education.

It should be made clear at the outset, however, that this Monograph has been prepared by and for educators. As a document it attempts to pull together insights and information gleaned from a number of disciplines with an eye toward a perspective upon which educational application can be built.

In the initial paper, the editors have portrayed comprehension as a setting prerequisite for and in which language processing occurs. This view challenges current conceptualizations of comprehension as outcomes or products or results, and it suggests that comprehension is not only process and product but also intent. The paper suggests that as researchers and teachers we need to become concerned not only with how meaning is constructed, but also with the decisions readers make on the way to the process—long before their eyes hit the page.

Tierney, Vaughn, and Bridge view comprehension from the point of view of inference. Their summation of recent approaches to the study of "comprehension as inference" is an important perspective, tying together several seemingly disparate attempts at studying and mapping the cognitive operations involved in discourse processing.

Tuinman views recent insights in comprehension from a measurement perspective and in so doing raises some concerns that he sees. Not least among these concerns is the notion that the lines drawn between general comprehension and reading comprehension are arbitrary and that focusing our attention on these lines may only constitute a distraction in learning to mean.

Page and Vacca systematically review current product and process measures in reading. They recommend the use of multiple measures and devise a corroborative framework for these measures.

Gould's study, as well as that by Rigg and Taylor, grows out of and supports insights into the measurement and process of comprehension presented by other authors in this volume. Each of these studies suggests that the semantic system of language is broad, that current measures of comprehension tap limited dimensions of the process, and that an expanded definition of comprehension is a prerequisite to pedagogical progress.

Smith's position—that we are essentially interested in how knowledge is acquired—is an important focus and a fundamental one when viewed from an educational perspective. The discussion narrows this focus by illuminating some old and new ideas that must be given credence if we are not to get sidetracked, lost, or both.

If Rhodes is right, many of the comprehending behaviors that children display are but simple artifacts of instruction. Rhodes' study demonstrates that the constraints that are set up and that operate in classrooms according to the kinds of instructional decisions which are made—in this instance, reading material selection—strongly affect cognitive and linguistic processing operations in reading. The study also suggests that the comprehending behaviors that are studied may be but aberrations of the "true" process.

New Perspectives on Comprehension is meant not only to organize perception and stimulate research, but also to suggest further instruction. In this latter regard we ought not only identify new instructional strategies but also weed through and abandon those that are either dysfunctional, misguided, counterproductive, or just plain egregious.

Tierney and Spiro identify comprehension principles that they perceive as being challenged by recent insights into the process. Because these principles should undergird instructional practice, their paper is but the beginning of an examination process in which each of us must engage if what has been learned is to result in change once the classroom and research doors have been closed.

New perspectives raise new concerns. Identification of these concerns, we believe, is part of the comprehending process. Each of the authors in this Monograph has raised educational concerns and has, in so doing, offered new perspectives to the intended audience.

We, too, have our concerns; for even as this book goes to print recent work by linguists and cognitive psychologists in collocational chains and metaphor have suggested yet other perspectives. We suspect there are still others that have escaped our attention.

Despite the advances in knowledge, we have much to learn, new gaps to fill, old shackles to abandon. What we now know is but a smidgen of what we will eventually come to understand. Yet, to delay this volume serves no one, for we can expect no less of ourselves than we do of readers; namely, to use what is now known and remain actively involved in the process of reading.

Goodman's paper bridges the gap not only between theory and practice, but between linguistics and politics. It too raises concerns, but is a fitting close as well as advent as its title, content, and message each singly captures best what this Monograph attempts to clarify.

New perspectives on comprehension define not only what comprehension is and is not, but also identify new variables which need to be studied. P. David Pearson's paper addresses some of the methodological concerns that have been raised relative to research on comprehension, offering not only new critique but also new conceptualization.

Jerome C. Harste
Robert F. Carey

Editors

Indiana University

Comprehension as Setting

Jerome C. Harste and Robert F. Carey

Jason and Heather, both 6, were eating lunch on the picnic table in Jason's back yard. For dessert, Jason's mother had fixed vanilla ice cream.

Heather, seeing Jason begin to mix his ice cream with his spoon, decided to do likewise.

After about a minute of stirring, Heather, with an Eureka-like tone in her voice, announced, "Oh Jason, I think we just found out how to make milk!"

A cute story illustrating childhood problems in comprehension, or is it? Let's examine what happened more closely.

For Heather to have made the statement she did, we know she had to have been busily monitoring not only the physical changes occurring in the dish of ice cream she was stirring, but relating these changes to other information which she already possessed from discovery of her world. It seems highly unlikely that anyone has specifically taught Heather that milk is white and liquid, while ice cream is white and in a semisolid frozen state; yet, quite obviously, she possesses this knowledge. From information available in the optic array, played against information she already has discovered about her world, Heather is able to make sense of this encounter by seeing a relationship which she had not seen before.

While we might like to argue, from our perspective, that the relationship which Heather saw was incorrect, a moment of reflection leads to the realization that in order for us to make this judgment, we had to do exactly what Heather had done, i.e., from information available in the optic array (Heather's statement as written in this article), we attempted to make sense (in this instance by reflecting on information we already think is known or shared—that milk is the key ingredient in making ice cream, not vice versa as Heather concluded).

From this perspective reading comprehension does not differ a great deal from comprehension generally. The process that Heather went through in arriving at her decision—that one made milk from ice cream—is highly similar, if not exactly like, the process we as readers went through in deciding she is wrong.

In this paper we wish to develop the thesis that comprehension is a setting for learning generally and for reading comprehension specifically. Conceptually this insight into reading comprehension is illustrated in Figure 1. This model proposes comprehension as a setting, involving both a print setting and a mental setting, with each in a sense providing an environment for the other. Given this conceptualization, the psychological realities of those insights into reading comprehension highlighted by this model will be examined and discussed.

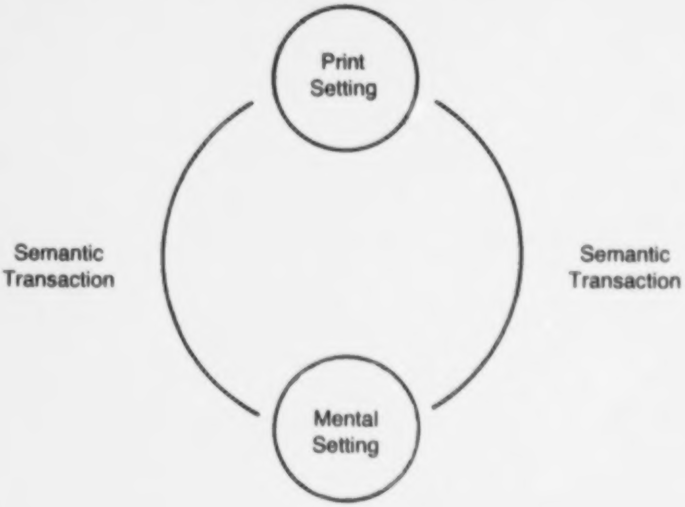


Figure 1. Comprehension as setting.

The Mental Setting

The circle entitled "Mental Setting" in Figure 1 is meant to accent the reader in the process of reading comprehension. For purposes of illustrating the importance of "mental setting" in comprehension, let us turn momentarily to Heather.

Heather, we can surmise, came to the setting not as a blank slate but with the accumulated wisdom of all she had discovered about her world. Her anticipation, as evident in her later surprise, is that ice cream will remain ice cream even when stirred on a nice summer day. The information available in the optic array allowed her not only to modify the schema or information she brought to the incident, but, further, permitted the discovery of a new relationship unseen before. Heather is not a passive recipient but an active participant in the process. To understand the dynamics of comprehension in this instance, one needs to be as concerned with Heather and what she did as with what is occurring with the ice cream. In actual fact the whole learning experience could have and would have been aborted if Heather had not remained cognitively active. It is Heather's confidence that things in her world make sense that allows her this opportunity for growth.

While it may be easy to see the important role that learners such as Heather play in making sense of their environment, the same has not been so, historically, in reading. For far too long the reader has played a stand-in role in studies of reading comprehension. By and large the spotlight in reading has remained on either the author or the text. Evidence of this fact is apparent when one examines tests that purport to measure reading comprehension. Here a template of the author's text serves as the touchstone from which judgments are made. While readers were not unimportant, their importance was secondary, used largely to measure, quantify, and norm information transfer. Precious few researchers have bothered to explore why a particular response was selected, and when they have done so the profession has tended to dismiss their efforts as descriptive rather than empirical.

"Mental Setting" as a concept is meant to reinstate and accent the importance of the reader if we are to understand his or her role in comprehension. Just as we gained insights into Heather's expectations through her expressed surprise, so too we must come to realize that the expectancies learners hold for alternate print settings provide valuable insights into reading comprehension.

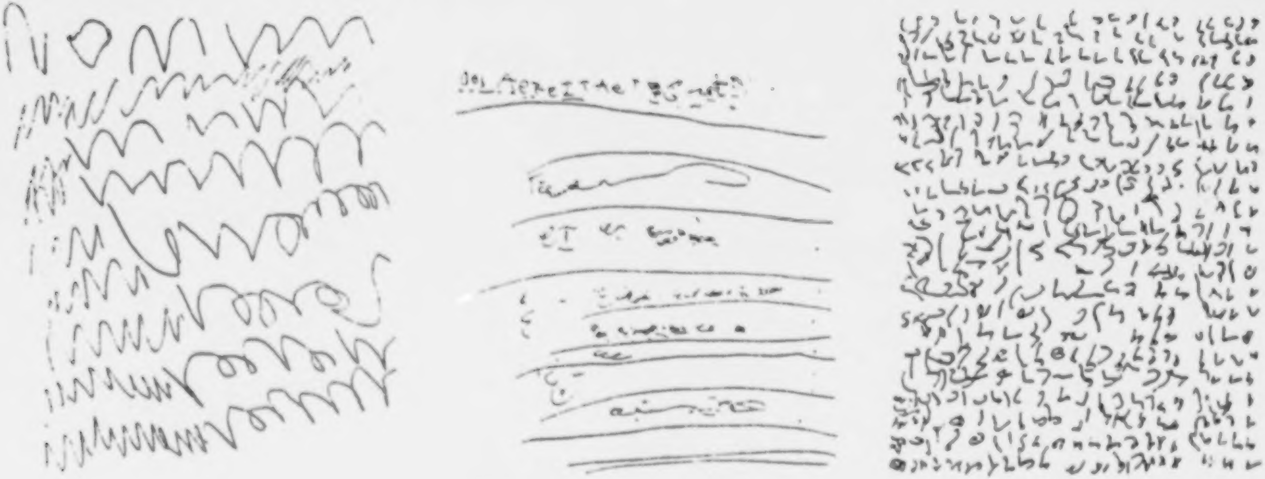
Probably no data more cogently documents the importance of print expectancies in understanding literacy than does that illustrated in Figure 2 collected from three 4-year-olds attending a preschool program. Uninterrupted samples of writing were collected from each of these children individually under the condition, "Write everything you can write."

Interestingly in contrast to the other two samples, Dawn's scribbles undeniably imitate English. When Najeeba finished her writing she said, "Here, but you can't read it because it is Arabic." Najeeba then went on to point out that in Arabic one uses "a lot more dots" than in English. Dalia is an Egyptian child whose "writing" bears an Egyptian look.

From ongoing natural encounters with print in each of their respective early language environments, these children have developed expectancies for print. What makes this data so exciting is that here is vivid evidence that (a) written language, like oral language, is learned; (b) children are actively involved in understanding and controlling their world, including the world of print; and (c) print expectancies are an important element within the acquisition of written language literacy.

Harste, Burke, and Woodward (in press) term the composite of what the child has discovered about language a "model," or series of interrelated strategies that allow the child to anticipate and respond predictably to encounters with alternate print settings. It is important to understand that they use the term model in its denotative sense, "as a series of assumptions through which experience is perceived and acted upon."

Further insight into the concept of mental setting and its relationship to reading comprehension is provided when young children are engaged in both uninterrupted reading and writing. A doctoral student, Hill (1978), asked 4-year-olds to write both a story and a letter. Typical of her data are the two samples produced



Dawn
United States

Najeeba
Saudi Arabia

Dalia
Egypt

Figure 2. Uninterrupted writing samples from three children age 4.

by Megan, Figure 3. While Megan's "writing" looks only slightly different, when asked to read what she had written, Megan's story and letter contrast sharply:

Megan's Letter (12-5-78): "Once upon a time—there was—wait—uh, huh. Dear Mary. I would like you to bring me. . . here everyday. The end. Megan."

Megan's Story (12-5-78): "Once upon a time there was a ghost. Three ghosts family. One day they went out for a walk. They honked the horn cause they saw Mrs. Wood and said 'I' then they went back to Mrs. Corners and they honked the horn and sa- said 'Hi.' The end."

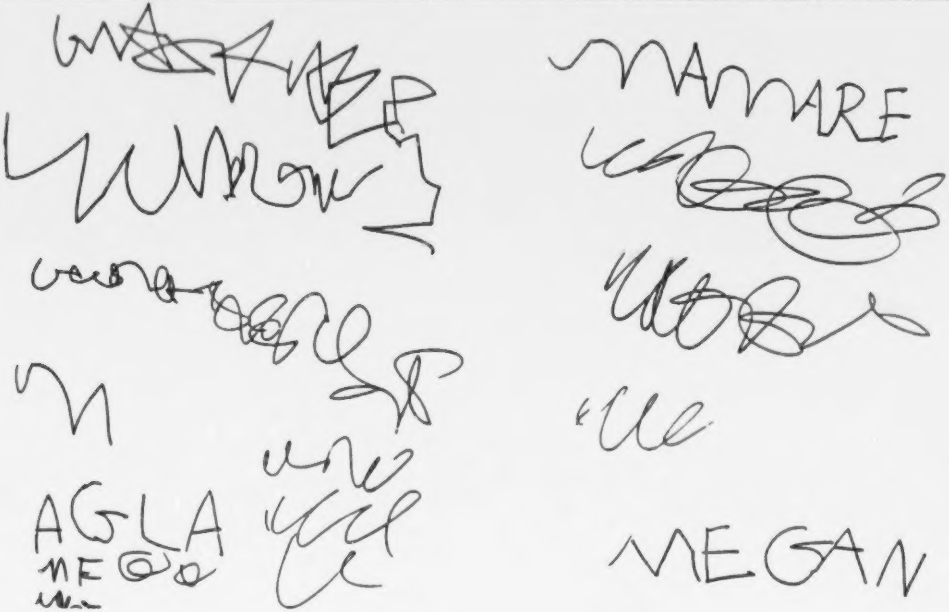
Not only does Megan's story have a story structure including a protagonist, attempt, and consequence (Stein, 1978), but her letter has a letter structure including salutation, body, and signature.

Megan demonstrates that from past encounters with print she has developed expectancies not only for the function but also for the form in which print will be cast. In Megan's writing one gets the clear notion that what she has discovered about alternate print settings is not isolated bits and pieces of language data, but a growing understanding of the fundamental structure that underlies these written language forms. Recent research by Stein (1978), Meyers (1977), and Kintsch and vanDijk (1978) in story grammar, story structure, and macrostructures, respectively, adds credence to these notions and begins to explore systematically how the reader's schema for alternate print forms affects comprehension and recall.

In a series of explorations in teaching (Barr, 1974-75; Harste & Burke, 1977, 1978; Harste, in press; DeFord, 1978; Mitchell, 1978) reader expectancies for print and strategy utilization during reading have been shown to be affected by the instructional model of reading being used in the classroom. Rhodes (beginning on p. 100 of this issue) discusses this factor more fully and demonstrates its effect upon print processing and comprehension. Other factors that have been shown to affect the reader's "mental setting" include background of experience and reader's purpose.

In a series of studies conducted at the Center for the Study of Reading, Anderson and his colleagues (Anderson, Reynolds, Schallert, & Goetz, 1977; Pichert & Anderson, 1977; Spiro, 1977) have cogently demonstrated that what the reader brings to the process strongly affects outcomes. In one study (Anderson, Reynolds, Schallert, & Goetz, 1977) ambiguous passages were given to a group of physical education students and a group of music education students. The findings of this study were predictable in that each group interpreted these passages as hypothesized in terms of their background of experience.

In an attempt to study how reader's purpose affects comprehension, Pichert and Anderson (1977) asked readers to read a passage from two perspectives—a potential homeowner and a burglar—only to find that quite different information was recalled as perspectives changed.



Megan
Story

12-5-78

Megan
Letter

12-5-78

Figure 3. Megan's story and letter: Uninterrupted writing samples.

At a more general and abstract level these examples of recent research illustrate the importance of "mental setting" in understanding reading comprehension. Restated formally, one's mental setting acts as anticipatory schema that direct sampling of information available in the print setting. From this perspective, the cognitive structures crucial for seeing are the anticipatory schema that prepare the perceiver to accept certain kinds of information rather than other kinds and thus control perception (Neisser, 1976). What cognitive structures one accesses in a given situation depends on experiences with a given print setting.

"Mental setting" as used here is a dynamic, though somewhat unmapped, concept. It is a construct that cannot be ignored, however, for if there is but one lesson to be learned from recent research in reading comprehension, it is that much of what is exciting in comprehension happens within the reader.

Uninterrupted writing samples and reading interviews can provide valuable insights into what expectancies readers hold for alternate print settings, just as purpose, instructional history, and uninterrupted oral reading samples can provide valuable insights into strategy utilization during reading and retelling. While admittedly the methodology at the moment is crude, the fact remains that we must first better understand the reader if we are to better understand reading and reading comprehension.

The position taken by our introduction of the concept of "mental setting" is that we must begin to honor this commitment by taking into account such factors as print expectancies, purpose, strategy utilization, background experience, and instructional history in future explorations into reading comprehension.

The Print Setting

Conceptually we have found it helpful to refer to readers as encountering "print settings" rather than as encountering print. This is in truth a matter of semantics. "Print setting" focuses attention on the notion of perception as robust, with any instance of print providing both linguistic and contextual cues. The context in which print is encountered provides many important cues (both visual and non-visual) for instantiation of schemata and thus comes to direct the sampling of available information in the optic array, the modification of anticipatory schema; in short, the process called cognition (Neisser, 1976; Smith, 1978). Perception and cognition, given this conceptualization, are not only the same, but both are dynamic.

The circle entitled "print setting" in Figure 1 is meant then to accent language in context and its importance to understanding comprehension generally and reading comprehension specifically. Just as Heather's statement, "Oh, Jason, I think we just found out how to make milk" has little humor in it when viewed as a statement outside of the specific context of situation, so, too, print takes on new meanings when found in context.

Figure 4 conceptualizes those dimensions of context included within the notion of pring setting. Conceptually this model hypothesizes that the context in which print is found affects both perception and cognition. Evidence in support of this idea is readily available when one examines recent research in psycholinguistics, sociolinguistics, and textlinguistics.

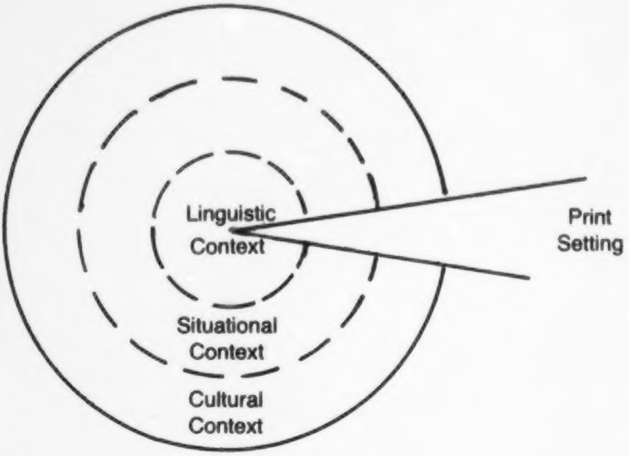


Figure 4. Print setting.

It has been found, for example, that individual letters become more perceptible when they are embedded in words (Reicher, 1968; Wheeler, 1970). Individual words are recognized more easily when they are embedded in meaningful sentences (Schuberth & Eimas, 1977; Tulving & Gold, 1963). Unknown words may be processed and even recognized when embedded in a familiar story context (Aukural, 1977; Goodman, 1965; Wittrock, Marks, & Doctorow, 1975). Longer sentences (regardless of their syntactic complexity, or whether in isolation or in discourse) that more coherently integrate underlying semantic relations are assimilated more easily than those that do not (Haviland & Clark, 1974; Pearson, 1976; Peck, 1977).

Recent works in propositional structures (Frederiksen, 1975; Frederiksen, Frederiksen, Humphrey, & Ottesen, 1978; Kintsch & vanDijk, 1978) and textual cohesion (Halliday & Hasan, 1976) have attempted to explore semantic frameworks within print and their relationship to natural language processing. To date these efforts have only begun to explore and specify what minimum cohesive ties are in text discourse from the perspective of readers, and what effect the absence of these ties have on reading performance (Cashen, Shelley, & Thomas, 1978; Rhodes, 1978; Tierney, 1978).

Context of situation has received even less study by the profession. Probably the most thorough explanation of this concept to date is that of Halliday (1974, 1975) via his concept of "register." For Halliday, "register" involves the three elements of *field*, *mode*, and *tenor*. In terms of discourse communication, *field* refers to the topic under discussion and specifically accents support information in the language environment that is available to assist content interpretation and processing. *Mode* refers to the expression of language used in the discourse and as such points to differences in interactions that result because of different expectancies for how speech as opposed to prose operates. *Tenor* refers to the relationship between language users in the discourse setting and suggests that the reader's relationship to others in conjunction with the print also affects perception and cognition.

The importance of context of situation in terms of understanding language processing is clear when one attempts to explain the divergent performance of students within particular studies. During an evaluation of the reading and writing abilities of Navajo students, Harste (1978) found, for example, that students had a more functional expectation for writing than for reading and this expectation strongly affected their performance in these areas. Harste and Burke (1978) have similarly found that better retellings can be obtained when readers are asked to retell what they have read to someone not present during their oral reading of the story. Harste, Burke, and Woodward (in press) report that first grade children respond quite differently to environmental print than do younger children after only 22 days of reading instruction in phonics.

These data begin to suggest that the constraints operative in one language setting are either different or not present in another language setting and that these constraints strongly affect the reader's perception and processing of print.

Context of culture, like context of situation, is an equally broad element within the notion of print setting. At one level, context of culture suggests that children growing up in literate societies develop alternate expectations for language based on the functions that language serves in that society. At another level, context of culture sees the classroom as a culture within which a particular form of print is found.

Probably the most interesting and insightful work in these areas to date has been done by sociolinguists. Bernstein (1971, 1973), Gumperz and Hymes (1972), and Cicourel (1974), for example, have explored how cultural norms are transmitted via language. Cazden (1965, 1978), Hymes (1967, 1978), Cazden, John, and Hymes (1972), McDermott (1977), and Greene & Wallat (1978) have studied the classroom as a cultural context and begun to explore how classroom cultural norms operate to affect language processing.

Taken as a conglomerate these studies speak to the importance of studying cognitive and linguistic processing in light of the varying contexts in which that process occurs. Print setting suggests the robustness of perception and implies that any instance of natural print provides multimodal clues which facilitate comprehension.

As a concept, "robustness of perception" is meant to portray perception as residing on a continuum of selection processes, the parameters of which are set by the reader's environment. Perception as robust pertains to processes other than reading; in fact, it seems pertinent to all processes involving sensory-based stimuli.

As Rosenblatt (1969, 1978) notes, both Dewey and Piaget, using varying terminology, have addressed a concept similar to robustness of perception. Piaget (1970) notes a process "whereby the organism in each of the interactions with the bodies of energies of its environment fits these [interactions] in some manner to the requirements of its own physico-chemical structures while at the same time accommodating itself to them" (p. 140). Dewey (1896) observes the same phenomenon, "something, not yet a stimulus. . . becomes a stimulus by virtue of relations it sustains to what is going on in this continuing activity. It becomes the stimulus in virtue of what the organism is already preoccupied with" (p. 255).

From this view, the reader, on the basis of information already possessed, previous encounters with print, and relevant aspects of the cultural, social, and situational contexts, anticipates and constantly monitors a variety of clues available for comprehension within the print setting. Additionally, a reader's perceptions of purpose and function can be viewed at least in part as the product of those aspects of local and global context that have a bearing on comprehension. "Perception as robust," then, is meant to portray perception not only as involving multimodal clues, but as a synergistic and cyclic process through which meaning is unfolded.

In a circuitous manner print setting accents the functional nature of written language. Another way, for example, to perceive this concept is to think of print as enveloped in discourse, discourse as enveloped in situation, and situation as enveloped in culture. Within any instance of print each of these contexts adds not only to the perceptability and understandability of print, but establishes it as functional. In many ways it is this dimension of functional quality, coupled with the robustness of perception, that makes written language predictable in terms of the anticipatory schema or mental setting that the reader brings to each new encounter with written language. In this sense the process is cyclic. What the reader learns from encountering a given print setting becomes the raw data from which to reformulate anticipatory schemata.

This concept of print setting is seemingly not well understood in most research in reading comprehension. In all too many explorations in reading, print is magically flashed in front of the eyes of subjects. While this may be "researching," it clearly is not reading. Print settings are predictable to readers who have had past encounters with print that were functional. To make print unpredictable is to study something that is not reading.

Semantic Transaction

Very recent research in reading has looked at reading as an interaction between author and reader (Adams & Collins, 1977; Frederiksen, et al., 1978; Kintsch & vanDijk, 1978; Rumelhart, 1976). Studies of this genre attempt to map retellings upon the text base in an effort to determine what contribution each participant—author and reader—makes. It was, in fact, our own frustrated attempts to study reading as an interaction that led to the conception of reading as semantic transaction.

Reading as semantic transaction, we now find, is not a new concept but a relatively old one, first introduced by Rosenblatt (1938). As a concept semantic transaction is meant to convey the idea that the product of reader meeting author is a totally new event.

Returning to Heather and Jason at the picnic table may help clarify what comprehension as semantic transaction means. Here Heather's proclamation, "Oh, Jason, I think we just found out how to make milk" can trace its origin neither to Heather's initial mental setting nor to the physical changes occurring in the dish of ice cream. Heather's response is clearly a new event—neither borrowed, mimicked, nor conditioned, but born out of the setting.

The same is true in reading comprehension, though probably less obviously so because of the constructs we have been using to study comprehension. A recent research experience illustrates the point.

In an effort to sort out the relative importance of background of experience from the context in which print is found, Harste, Carey, and Smith (in preparation) replicated a study by Anderson et al. (1976) in which they confounded these variables, but concluded that the schema which readers bring to reading strongly affects their interpretation of print.

Specifically Anderson et al. gave groups of physical education students and music education students the following ambiguous passages:

Passage 1—Prison/Wrestling[1]

> Louie slowly got up from the mat, planning his escape. He hesitated a moment and thought. Things were not going well. What bothered him most was being held, especially since the charge against him had been weak. He considered his present situation. The lock that held him was strong but he thought he could break it. He knew, however, that his timing would have to be perfect. Louie was aware that it was because of his early roughness that he had been penalized so severely—much too severely from his point of view. The situation was becoming frustrating; the pressure had been grinding on him for too long. He was being ridden unmercifully. Louie was getting angry now. He felt he was ready to make his move. He knew that his success or failure would depend on what he did in the next few seconds.

Passage 2—Music/Cards

> Every Saturday night, four good friends get together. When Jerry, Mike, and Pat arrived, Karen was sitting in her living room writing some notes. She quickly gathered the cards and stood up to greet her friends at the door. They followed her into the living room but as usual they couldn't agree

---

[1] Between the time of Anderson's study and the replication described here, the Academy Award winning movie "Rocky" was released nationwide. Because we were studying the relative effects of background of experience and context, the character's name in the passage was changed to Louie instead of Rocky to avoid biasing the data toward this contextual reference and hence influence interpretation.

> on exactly what to play. Jerry eventually took a stand and set things up. Finally, they began to play. Karen's recorder filled the room with soft and pleasant music. Early in the evening, Mike noticed Pat's hand and the many diamonds. As the night progressed the tempo of play increased. Finally, a lull in the activities occurred. Taking advantage of this, Jerry pondered the arrangement in front of him. Mike interrupted Jerry's reverie and said, "Let's hear the score." They listened carefully and commented on their performance. When the comments were all heard, exhausted but happy, Karen's friends went home.

Each passage, according to Anderson et al. could be interpreted in one of two ways: The first passage as either a wrestler attempting to break a hold or as a person planning an escape from prison; the second passage as either an ensemble group coming together to practice or as a group of people coming together to play cards.

Because we were interested in showing that Anderson et al. confounded reader background experience with print setting (arguing the music education students in a music education class, for example, had every right to expect print related to music) we not only replicated the original study, but further attempted to collect data from music majors in settings (e.g., in English classes) where their expectation for musical print would not be present. Similarly, we attempted to collect interpretations from groups of physical education majors as we increased the contextual clues available (from classroom to sitting on mats in gyms, for example) to study the effects of context on comprehension.

Upon completing our first round of data collection (essentially an effort to replicate the original study) we sat down to code written retellings by major theme: playing music/playing cards or wrestling/prison. The coding of themes, which was projected as a one-night task by each of us so that we could determine inter-rater reliability, proved excruciatingly complex. Themes identified from the written retellings of the music/card playing passage, rather than falling into two nice categories, ranged from playing music *to* listening to music *to* playing games *to* playing cards *to* talking *to* having sexual relationships *to* combinations of these themes being the hallmark of any given retelling rather than the exception.

Several examples illustrate some of the more interesting, yet typical, problems we encountered.

Example 1

> This passage is about four friends who get together on a Saturday night. There are two guys and two girls. They play cards for a while and listen to music, after a while things begin to get dull and one of the males begins to gain interest in one of the females. Then they begin to notice their sexual difference and study the "arrangement."
>
> Research Notes
>
> Themes identified include card playing, listening to music, sex.
> Ambiguity of original text noted via quotation marks.

Example 2

> Karen and three of her friends had gathered at Karen's house to play some card game. After not being able to decide what game to play, a table was set up by Jerry, and they played for a while to music from the tape recorder. After a while, when play got more intense, one of the guys asked what the score was. We never were told what it was, but soon the group discussed how they did that night, then broke up and went home.
>
> Research Notes
>
> Themes identified include card playing and listening to music from the tape recorder.
> Ambiguity of original text noted by reference to the fact that scores were never given; reference would suggest reader expected this information after author introduced topic.

Example 3

> Every Saturday night, four friends get together. When Jerry, Mike, and Pat got to Karen's house Karen was writing notes in the living room. She put away her cards and went to greet her friends. They couldn't decide what to play until finally Jerry took charge of the situation. Karen played the recorder beautifully. Mike noticed the beautiful diamonds on Pat's hand. After they finished playing everyone went home.
>
> Research Notes
>
> Effort to retain original ambiguity: Karen plays music with ambiguous reference to the group having played something.

Even students with music majors interpreting the music/card playing passage as a group of musicians coming together expressed puzzlement over phraseology by writing, "You don't 'listen' to a score." Others who saw the passage as having a card playing theme nonetheless saw Pat's diamonds as a ring and not as a suit of cards.

The wrestling/prison passage caused no fewer problems in coding. Themes identified from written retellings collected from students in our sample ranged from a wrestler breaking a hold *to* an attempted prison break *to* a hostage being held *to* a child being punished *to* an animal being tied up *to* any number of permutations on these themes.

Part of this rich variety in responses can be captured through comparison of a random sample of 10 opening lines in our retelling data (N = 142) collected from the wrestler/prison passage:

Wrestling Theme:

> Louie is a wrestler.
> Louie was in some form of activity and was planning to escape.
> The passage deals with the thoughts and decisions going through Louie's head as he sizes up the predicament he has found himself in on the wrestling mat.

Prison Theme:

> Louie was in jail and he got up from a mat he was lying upon while he was planning his escape.
> The passage described one man's mental journey of escaping from jail.

Other Themes:

> Louie is imprisoned, in his mind, being held unfairly with little type of evidence.
> Louie is being detained in chains on a mat.
> Louie has been taken "hostage" by someone for someone for some reason.
> Louie is being "punished" for something he felt was unjust.
> Louie is a kid who is punished by his parents for something that he did (bad).

What becomes readily apparent from an examination of these responses is not their similarity, but their difference even within the same theme. From visual cues available each of these readers was able to create a sense of situation. What occurred was not a poor rendition of the text, but a unique event—in essence a new text as original and distinctive as the author's.

Reading as semantic transaction is meant to capture this dynamic change. Comprehension in each of these instances is best seen as an active search for meaning. The reader, rather than being duped by the mental setting which was brought, seems to have been particularly sensitive to the comprehension clues provided by the print setting. The result is a fluid, give-and-take between mental setting and print setting in an effort to make sense of the story.

Many of the readers in our study were particularly bothered by the ambiguous passages used in this study and indicated such in their retellings. In a variety of ways, students seemed to be trying to express their feeling that some implicit contract between author and reader had been broken. Tierney (1978), too, has noted this phenomena and calls it an "implicit allowability contract" between author and reader.

Other readers attempted to write their retelling so that it retained the author's original ambiguity. Often the results were surprisingly close to the original, despite interpolated tasks that would clearly prohibit retention of the passage as sensory image, as indicated by the retellings below:

Example 1

> Karen had her friends over again Saturday night as usual. She was copying her notes and put them up when the friends arrived. No one said what they wanted to play so Jerry decided and set things up. They all played and when the fun was gone they quit, announced the scores and, exhausted, the friends went home.

Example 2

> Four good friends, Jerry, Mike, Pat, and Karen meet together. As usual they did not agree on what to play when they arrived at Karen's place. Finally they began to play with Karen's recorder playing soft music. As the night progressed the tempo increased but during a lull Mike wanted to hear the score. After commenting they went home.

Example 3

> On every Saturday evening, four friends get together. Karen was sitting at home, and her three friends came to the door. She put down the cards and let them in. They couldn't decide what to play, but one set things up and they began to play. Karen's recorder filled the room with a beautiful pleasant sound. One boy noticed diamonds on Pat's hand. The three friends of Karen's left the house tired but happy.

Example 4

> Four friends get together every Saturday night. Jerry, Mike, and Pat arrived at Karen's house that night. Karen was working on some notes when they arrived. She went to greet them. They didn't know what to play. Jerry got a stand and set things up. Soon the music from Karen's recorder was filling the air. Jerry noticed a diamond ring on Pat's hand while they were playing. Soon it was time to leave.

One cannot help but wonder, when examining these retellings, what relationship past encounters with print in a classroom test situation have had in shaping this kind of behavior. Clearly this retelling behavior seems unnatural, but may serve to accent the dynamic interrelationship between print, print settings, print expectancies, perception as robust, mental setting, and reading comprehension.

Conclusion

This paper has attempted to (1) destroy the myths surrounding the process of reading comprehension by tying it more closely to other quite natural human activities like thinking and learning; (2) propose comprehension as a condition for language processing, thus opening up the possibility of looking at reading comprehension not only as product and process but intent; (3) reinstate the importance of the reader in studies of reading comprehension by identifying key factors and relationships that impinge upon the reader's model of reading and which affect cognitive and linguistic processing; (4) establish the importance of studying reading comprehension within those ecologically valid contexts in which that processing naturally occurs; and (5) cast reading comprehension as a dynamic give-and-take between author and reader.

While insights into reading comprehension offered by such concepts as mental setting, print setting, perception as robust, comprehension as semantic transaction, and comprehension as setting undoubtedly complicate the task of the researcher, seeing these relationships, we believe, opens the vistas for new conceptualizations and much needed explorations in written language literacy and reading comprehension.

References

Adams, M. J., & Collins, A. *A schema-theoretic view of reading* (Tech. Rep. No. 32). Urbana-Champaign, Ill.: University of Illinois, Center for the Study of Reading, 1977.

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. *Framework for comprehending discourse* (Tech. Rep. No. 12). Urbana-Champaign, Ill.: University of Illinois, Laboratory for Cognitive Studies in Education, 1976. Also published in *American Educational Research Journal*, 1977, *14*, 367-381.

Aukural, K. R. *The epistemological assumptions of the Goodman-Burke immediate context strategy and a case study of its efficacy.* Unpublished doctoral research paper, Reading Department, Indiana University, 1977.

Barr, R. The effect of instruction on pupil reading strategies. *Reading Research Quarterly*, 1974-75, *10*, 555-582.

Berstein, B. *Class, codes, and control I: Theoretical studies towards a sociology of language*. London: Routledge & Kegan Paul (Primary Socialization, Language, & Education), 1971.

Berstein, B. (Ed.). *Class, codes, and control II: Applied studies towards a sociology of language*. London: Routledge & Kegan Paul (Primary Socialization, Language, & Education), 1973.

Cashen, C. J., Shelley, A., & Thomas, S. *Inferencing required by basal first grade readers: A look at phonics, skills, and the whole language approach to first grade reading and inferences required by each.* Unpublished doctoral research, Reading Department, Indiana University, 1978.

Cazden, C. B. *Environmental assistance to the child's acquisition of grammar.* Unpublished doctoral dissertation, Harvard University, 1965.

Cazden, C. B. *Toward educational linguistics.* Paper presented at the Annual Meeting of the National Council of Teachers of English, Kansas City, 1978.

Cazden, C., John, V., & Hymes, D. (Eds.). *Function of language in the classroom.* Teachers College Press, 1972.

Cicourel, A. V. *Cognitive sociology.* Free Press, 1974.

DeFord, D. E. *A validation study of an instrument to determine a teacher's theoretical orientation to reading instruction.* Unpublished doctoral dissertation, Indiana University, 1978.

Dewey, J. The reflex arc concept in psychology. *Psychological Review*, 1896, *3*.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458.

Frederiksen, C. H., Frederiksen, J. D., Humphrey, F. M., & Ottesen, J. *Discourse inference: Adapting to the inferential demands of school texts.* Paper presented at the Annual Meeting of the American Educational Research Association, Toronto, Canada, 1978.

Goodman, K. S. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, *42*, 639-643.

Greene, J., & Wallat, C. *Sociolinguistic ethnography: A methodology for identification of social interaction contexts*. Paper presented at the Annual Meeting of the International Reading Association, Houston, May 1978.

Gumperz, J. J., & Hymes, D. (Eds.). *Directions in sociolinguistics: The ethnography of communication*. New York: Holt, Rinehart, & Winston, 1972.

Halliday, M. A. K. *Language and social man*. London: Longman, 1974.

Halliday, M. A. K. *Learning how to mean*. Edward Arnold Ltd., 1975.

Halliday, M. A. K., & Hasan, R. *Cohesion in English*. London: Longman, 1976.

Harste, J. C. Reading and reality. In J. Robinson (Ed.), *Whole language for whole children*. Columbia, Mo.: Reading Department, University of Missouri, in press.

Harste, J. C. *Evaluation report: Wingate Language Arts Program*. Tuscon: Center for the Expansion of Language and Thinking, 1978.

Harste, J. C., & Burke, C. L. A new hypothesis for reading teacher education research: Both the teaching and learning of reading are theoretically based. In P. D. Pearson (Ed.), *Reading: Research theory and practice*. 26th yearbook of the National Reading Conference, Mason Publishing Co., 1977.

Harste, J. C., & Burke, C. L. *Retellings as cognitive processing*. Paper presented at the Annual Meeting of the National Reading Conference, St. Petersburg, December 1978.

Harste, J. C., & Burke, C. L. Toward a socio-psycholinguistic model of reading comprehension. *Viewpoints in Teaching and Learning*, 1978, *54*(3), 9-34.

Harste, J. C., Burke, C. L., & Woodward, V. Children, their language and world: Initial encounters with print. In M. Smith-Burke & J. Langer (Eds.), *Author meets reader*. Newark, Del.: International Reading Association, in press.

Harste, J. C., Carey, R. F., & Smith, S. L. *Frameworks for understanding comprehension: A replication*. (In preparation)

Haviland, S. E., & Clark, H. H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 512-521.

Hill, M. W. *Look I can write: Children's print awareness from a socio-psycholinguistic perspective*. Unpublished doctoral research, Reading Department, Indiana University, 1978.

Hymes, D. *Theory, practice, and prestige*. Paper presented at the Annual Meeting of the National Council of Teachers of English, Kansas City, 1978.

Hymes, D. H. Models of interaction of language and social setting. *Journal of Social Issues*, 1967, *23*.

Kintsch, W., & vanDijk, T. A. *Toward a model of text comprehension and production*. Unpublished paper, Department of Psychology, University of Colorado, Boulder, May 1978.

McDermott, R. P. The ethnography of speaking and reading. In R. Sharp (Ed.), *Linguistic theory*. Newark, Del.: International Reading Association, 1977.

Meyers, B. J. F. What is remembered from prose: A function of passage structure. In R. O. Freedle (Ed.), *Discourse production and comprehension*. Ablex, 1977.

Mitchell, K. *Patterns of teacher-student interactions during reading as related to teachers' theoretical positions*. Paper presented at the Annual Meeting of the National Council of Teachers of English, Kansas City, 1978.

Neisser, U. *Cognition and reality*. Freeman, 1976.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relation. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*. Newark, Del.: International Reading Association, 1976.

Peck, A. *Relations between syntactic complexity in connected discourse and children's ability to make inference*. Unpublished doctoral research, Reading Department, Indiana University, 1977.

Piaget, J. *Structuralism*. Chaninah Maschler (trans.) New York: Basic Books, 1970.

Pichert, J. W., & Anderson, R. C. Taking different perspectives on a story. *Journal of Educational Psychology*, 1977, *63*(4), 309-315.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus materials. *Journal of Experimental Psychology*, 1968, *81*, 275-280.

Rhodes, L. K. *The interaction of beginning readers' strategies and texts reflecting alternate models of predictability*. Unpublished doctoral dissertation, Indiana University, 1978.

Rosenblatt, L. M. *Literature as exploration*. New York: Appleton-Century, 1938.

Rosenblatt, L. Towards a transactional theory of reading. *Journal of Reading Behavior*, 1969.

Rosenblatt, L. M. *The reader, the text, the poem*. Carbondale, Ill.: Southern Illinois University Press, 1978.

Rumelhart, D. E. *Toward an interactive model of reading* (Tech. Rep. No. 56). San Diego: Center for Human Information Processing, University of California, 1976.

Schuberth, R. E., & Eimas, P. D. Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 27-36.

Smith, F. *Understanding reading* (2nd ed.). New York: Holt, Rinehart, & Winston, 1978.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Erlbaum, 1977.

Stein, N. L. *How children understand stories: A developmental analysis* (Tech. Rep. No. 69). Urbana-Champaign, Ill.: Center for the Study of Reading, University of Illinois, March 1978.

Tierney, R. J. *The discourse processing operations of children.* Paper presented at the Annual Meeting of the National Reading Conference, St. Petersburg, December 1978.

Tulving, E., & Gold, C. Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology,* 1963, *66,* 319-327.

Wheeler, D. C. Processes in word recognition. *Cognitive Psychology,* 1970, *1,* 59-85.

Wittrock, M. C., Marks, C., & Doctorow, M. Reading as generative process. *Journal of Educational Psychology,* 1975, *67,* 484-489.

Jerome C. Harste, Ph.D., is associate professor of education, Indiana University, Bloomington. Robert F. Carey, Ph.D., is assistant professor of education, Indiana University, Bloomington.

Toward Understanding Comprehension: An Examination of Systems for Analyzing Inferences

Robert J. Tierney, Joseph L. Vaughan, Jr.,
and Connie A. Bridge

In recent years, a number of frameworks have been offered for the examination of inferences by researchers and theorists intent on addressing the diverse nature of inference based on cognitive concerns and sociolinguistic parameters. A basic postulate of these frameworks is the notion that even from a "text-based" understanding of what is read, reading comprehension involves reader inference rather than explicit text-based understandings. As Kintsch (in press) has suggested, "Meaning is something the reader creates in response to a text, not something directly given by the text" (p. 14).

In regard to suggesting parameters for any examination of inference, the present paper examines three frameworks for studying inferences and addresses their adequacy. The three frameworks are a Progressive Refinement Model (Collins, Brown & Larkin, 1977), Trabasso's Inferences for a Narrative (Trabasso & Nicholas, in press; Warren, Nicholas, & Trabasso, in press), and Frederiksen's Examination of Text-Based Inferences (Frederiksen, 1977a, 1977b; Frederiksen, Frederiksen, Humphrey, & Ottesen, 1978).

Progressive Refinement Model

Many contemporary theorists believe that reading comprehension occurs and proceeds as a result of inferences the reader makes during discourse processing. These inferences, they suggest, are guided by the reader's schemata, i.e., active mental structures that guide the reader's pursuit and derivation of meaning during discourse processing. In other words, it is within the reader's schemata that the basis for construction of meaning through inferencing exists, and it is from within those schemata that the interaction of reader with text is guided to this end (Anderson, 1977; Rumelhart & Ortony, 1977; Spiro, 1977).

Collins et al. (1977) have suggested a framework for examining inferences based on these schema-theory notions. A basic postulate of their framework is that a reader's inferences are guided by the progressive refinement of a scenario or model that the reader develops for a text. They suggest that text understanding proceeds and inferencing occurs via the refinement of the reader's model.

> The initial model is a partial model, constructed from schemas triggered by the beginning elements of the text. The models are progressively refined by trying to fill the unspecified slots in each model as it is constructed. . . and the search for relevant information is constrained more and more. (pp. 4-5)

Within this framework, they suggest that inference is involved in the structuring, restructuring, presupposing, binding, rebinding, focusing, assigning, and overall constraining of variables. In other words, it is by inference that the reader refines a model or scenario which "matches" the text and which is complete, interconnected, and plausible. Thus, according to Collins et al., inference will play a number of different roles toward the progressive refinement of a particular model for a text. For example, inference will be involved in the reader's construction of a scenario to account for the elements and relationships within the text and the world as the reader sees it. If the reader's model seems tenable, then inference will be involved in the further elaboration and binding of elements and relationships. If the reader's model seems untenable, then inference will be involved in the re-examining, reconstructing, or restructuring of elements and their relationships.

To illustrate these processes, consider what is involved in understanding the following text (Collins et al., 1977).

Window Text

> He plunked down $5 at the window. She tried to give him $2.50, but he refused to take it. So when they got inside, she bought him a large bag of popcorn. (p. 3)

With the initial statement, "He plunked down $5 at the window," readers often generate scenarios involving a racetrack window, bank window, or a theater. They address pending questions, such as "who?" "where?" and "why?" As depicted in Figure 1, they may have constructed, for example, an initial model of the text involving a bank window with some bound variables (concepts enclosed by boxes) and some variables awaiting binding (concepts enclosed by circles). As the reader proceeds, these scenarios are progressively refined: "She" is usually defined as the recipient of the money; "$2.50" is usually identified as change. Then with the statement, "So when they got inside. . . popcorn," readers usually recognize a conflict. They realize their models no longer match the text and are implausible, disconnected, and incomplete. To restructure their model, they might question previous interpretations, for example, the female was a bank clerk or a ticket vendor, and maybe shift their focus, for example, from a bank to a theater. Eventually, it is likely that a scenario will evolve for the reader that involves the purchase of two tickets and an attempt by a date to share the expenses.

In accordance with schema theory notions then, the proposal of Collins et al. provides parameters for examining inference that enable an analysis of the nature of the constructive interaction between text and reader. That is to say, their proposal entails examining inference holistically, within the framework of the reader's construction, evaluation, and perhaps revision of a scenario for the text.

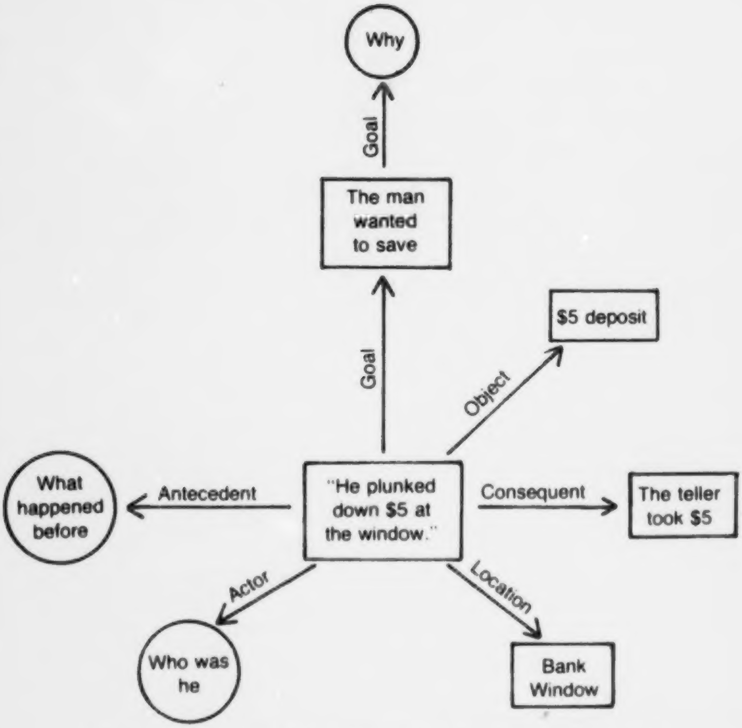

Why
Goal
The man wanted to save
$5 deposit
Goal
Object
What happened before
Antecedent
"He plunked down $5 at the window."
Consequent
The teller took $5
Actor
Location
Who was he
Bank Window

Figure 1. Initial stage in construction of model for window text.

Trabasso's Inferences for a Narrative

Trabasso and Nicholas (in press) and Warren et al. (in press) view understanding as involving the representation of incoming textual information in a well-defined structure. To this end, they suggest, "inferences serve two main functions: they (1) fill in missing slots in the structure, and (2) connect events with other events in order to provide a higher level of organization" (Warren et al., in press). Toward modeling the process of making inferences, Trabasso and Nicholas as well as Warren et al. restrict their examination of inference to the context of narratives based on event chain formulations.

An event chain formulation represents what happened in a given story. The links of the event chain correspond to the single episodes or actions explicitly stated within the narrative in the order they occurred. In its formal representation, an event chain depicts, by protagonist, several broad classes of events (states, events, actions, cognitions, displays, impulses, and goals) and logical connectives (motivation, physical cause, psychological cause, enablement, temporal succession, and temporal coexistence). Certain a priori rules constrain the possible combination of propositional types and connectives. For example, only certain classes of events (action, display, and event) can have a causal relationship that is physical. In stories involving multiple protagonists, the events related to each protagonist shift horizontally in accordance with a shift in characters.

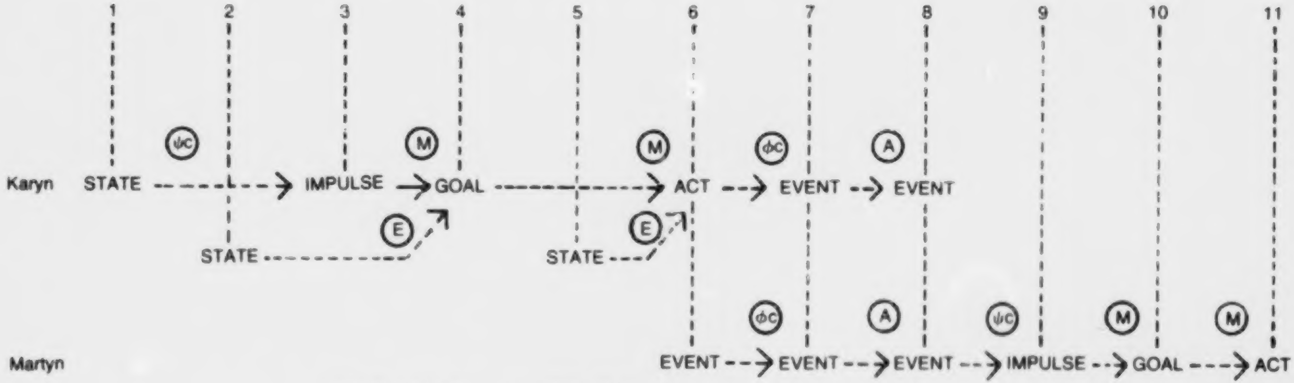
As an illustration of the structure of an event chain, consider the following brief story and its depiction in Figure 2. The figure depicts the event chain of the story with each event numbered and labeled. Their interconnections are represented by labeled arrows; the shift in protagonist is depicted by a shift in horizontal lines.

1. It was the weekend.
2. Martyn was playing in the sand tray.
3. Karyn felt mischievous.
4. She decided to tease Martyn.
5. When Martyn was not looking,
6. she turned the hose on.
7. Martyn was covered by water.
8. He grabbed the hose.
9. He was very angry.
10. So to get even with Karyn,
11. he sprayed her.

In conjunction with their formulation of event chain, Trabasso and Nicholas (in press) and Warren et al. (in press) propose a taxonomy of inferences. Their taxonomy provides categories for the types of inferences a reader might make within and across event chains. It is interesting to note that in the authors' views their categories of inferences are seen as neither discrete nor hierarchical. As Trabasso and Nicholas point out "No class of inference can operate *in vacuo*, and there is no fixed hierarchy of classes."

The categories of inference within the taxonomy include three broad types: informational, logical, and evaluative. The informational inferences involve the determination of the "who," "what," "when," and "where" within stories. They include pronominal inferences, referential inferences, spatio-temporal inferences, world frame inferences, and elaborative inferences. The logical inference category addresses the "how" and "why" of stories. This category includes inferring motivation, psychological cause, physical cause, and enablement. Value inferences address the "so what" of the story. They include the reader's evaluation of the events, characters, construction, and style of the story. The inferences within each of these classes are presented in Table 1.

In an attempt to define practical limits to inferencing, Trabasso and Nicholas (in press) and Warren et al. (in press) address what they term a "relevancy hypothesis." The relevancy hypothesis states that the reader, understanding a narrative, should make only those inferences determined by and integral to the progress of the narrative. In other words, they suggest that the reader should make only those inferences necessary to determine what happened and why.


1
2
3
4
5
6
7
8
9
10
11
Karyn
STATE
ψc
IMPULSE
M
GOAL
M
ACT
φc
EVENT
A
EVENT
STATE
E
STATE
E
φc
A
ψc
M
M
Martyn
EVENT
EVENT
EVENT
IMPULSE
GOAL
ACT

Key to symbols	
	φc —physical cause
	M —motivation
	ψc —psychological cause
	E —enablement
	T —temporal succession
	A —temporal coexistence

Figure 2. Event chain for the narrative involving Karyn and Martyn.

Table 1
Taxonomy of Inferences During Story Comprehension

Classes	Functions
1. Logical Inference	
a. Motivation	Inferring causes for a character's given voluntary thoughts, actions, or goals (or vice versa), e.g., John was angry. He left.
b. Psychological cause	Inferring causes for a character's given involuntary thoughts, actions, or feelings (or vice versa), e.g., John tripped on the stone. He shouted.
c. Physical Cause	Inferring mechanical causes for given objective events or states (or vice versa), e.g., Lightening hit. The tree fell. David smashed the car. His passenger was injured.
d. Enablement	Determining the conditions necessary but not sufficient for a given event to occur. Determine the event a certain condition allows, e.g., It was windy. They could fly the kite.
2. Informational Inferences	
a. Pronominal	Specify the antecedents or pronouns, e.g., Chuck was late. He was mad.
b. Referential	Specify the related antecedents of given actions or events when the reference is not pronominally marked, whether or not they are explicitly stated in other propositions, e.g., Carol found her father's car in front of the school. She ran and hopped in.
c. Spatio-Temporal	Determine the place or time of a single or series of propositions, e.g., It was Friday afternoon. They ran to the football park. The children were all ready.
d. World-Frame	Determining a world context to account for inferences, e.g., They saw the lions, tigers, seals, and monkeys.

Table 1 con't

e. Elaborative	Flushing out additions which do not contribute to the logical process of the story.
3. Value Inferences	
a. Evaluative	Judging the morality, convention, and anomaly in character's thoughts and actions or in story style or construction, e.g., John wanted to tease Peter. He asked him if he could leave the party. Peter was shocked and angrily shouted obscenities. Was shouting obscenities a good way to deal with John?

---

[a] Based on Warren et al. (in press).

Alternatively, they suggest that certain inferences may be superfluous or redundant. That is, while these inferences may be consistent with the text and add color to the story, they are not relevant to the flow of the narrative. Indeed, elaborations, spatio-temporal inferences, referrents, and enablements may often be deemed irrelevant.

In essence, then, according to Trabasso and Nicholas (in press) and Warren et al. (in press), a reader makes progress within the context of a narrative by engaging two inferential processes: slot filling and connecting text information. These processes are, or should be, directed toward following "what happened" and "why." According to the taxonomy this might include the "who," "what," "why," "how," and "so what" of a story.

Frederiksen's Examination of Text-Based Inferences

Frederiksen (1975) views understanding or comprehension "as a process whereby a listener or reader attempts to infer the knowledge structure of a speaker or writer by using the available linguistic message, contextual information, and his or her own knowledge store as data structures from which the inference is to be made" (p. 371). Based on the premise that an examination of inference must account for the interplay between text-based and knowledge-based processes, Frederiksen (1975, 1977a, 1977b) offers a framework for examining inference that purports to address the text, reader, and communicative context (Frederiksen et al., 1978).

To illustrate and explain the use of Frederiksen's framework, consider a reader's recall of the following sentences. The sentences were taken from a story entitled

"Johnny and the Squirrel"; the recall was taken from a reader's recall for the entire selection.

> His shirt was jumping back and forth. His mother came running.

Recall:

> The shirt was jumping back and forth on the bed. Then Johnny's mother ame running.

At the lowest level of analysis, Frederiksen's framework would define the semantic content and logical structure of the text.[1] In brief, the semantic content consists of propositions that are represented as networks of concepts connected by labeled binary relations. The logical structure represents the logical, causal, and algebraic relations between propositions. For example, the semantic content and logical structure of the text above might be represented by the numbered propositions in Figure 3. The abbreviated symbols denote some of the concepts and relationships defined by Frederiksen's semantic and logical network system.

To complete the first level of analysis, a similar representation would be derived from the reader's recall, then both representations would be compared. Based on this comparison, the second and third levels of analysis proceed. These analyses entail (a) an examination of the explicit/inferred information within a reader's recall, and (b) a taxonomical examination of and functional analysis of inferential operations. Specifically, these analyses would entail the identification of the different types of inferences made by a reader and the functions these inferences served. For example, in the previous recall, the reader engaged an ordinal connection "then" between the first and fourth propositions. That is, the reader engaged a relative operation toward segmenting the text into sets of interrelated propositions within the event structure of a story (the function of the inference). Likewise, in the second sentence the reader's substitution of "Johnny's" for "his" suggests resolution of anaphora; in the first sentence the reader's insertion of "on the bed" suggests an inference serving a locative or deitic function. Table 2 includes a listing of Frederiksen's taxonomy of text-based inferences; Table 3 includes a listing of the functional contexts of inferences, and Figure 4 shows their contribution to discourse comprehension.

At these second and third levels, therefore, Frederiksen's framework affords a concurrent text-based analysis of inferences and a functional examination of their role. Unfortunately, many researchers working with this system fail to extend their examinations to address these levels of analysis.

---

[1] Frederiksen's model grows from an attempt to extend the sentence-based semantic network proposed by Simmons (1972). It incorporates a taxonomy of objects and actions similar to that suggested by Chafe (1970); specification of causal circumstances and participants in events based on Fillmore's (1968, 1971) work on case categories and Chafe's (1970) notions relative to the relation of nouns to verbs; dimensional structure and treatment of tense and aspect based on Leech (1969); a notational system based on Simmons (1972) and Schank (1972, 1973); psychological measurement theory pertaining to metric data structures; and, the notions of Rescher (1964) and Simon and Rescher (1966) relative to "cause and counterfactual."

01 ('Johnny) - PAT @ TEM (PRES) - (has) - OBJ - (:shirt)

(shirt) - DEF - NUM - (one)

02 ("02) - AGT @ TEM (PAST) ASPECT (CONT) - (jump) - MAN - (back and forth)

03 ('Johnny) - PAT @ TEM (PAST) - (has) - DAT - (:mother)

04 ("04) - AGT @ TEM (PAST) - (came) - MAN - (running)

Key to Symbols in Network

( )	concept
(:)	concept to be determined and quantified
(')	concept not to be determined and quantified
(")	reference to proposition usually cited by number
@	marks an operation on the relation

Relations

Case relations, resultive propositions

AGT	Agent	participant in the act
DAT	Dative	recipient of the act (animate)
OBJ	Object	recipient of the act (inanimate)

Case relations, processive propositions

PAT	Patient	participant in the act
DAT	Dative	recipient of the act (animate)
OBJ	Object	recipient of the act (inanimate)

Other

MAN	Manner adverbial
DEF	Determination (definite)
TOK	Determination (indefinite)
NUM	Quantification
PAST	one of a variety of tenses
CONT	one of a variety of aspects

Figure 3. Message base (knowledge structure) of text.

Table 2
Frederiksen's Taxonomy of Text-Based Inferences

Class	Definition
I. Identification Relations	
1. Lexical Operations	Determine what information is lexicalized and what information is expanded into propositions
2. Object Identification	Specifies how an object or set of objects is related to or distinguished from other objects.
3. Action Identification	Specifies how an action is related to or distinguished from other actions.
4. Locative Identification	Specifies a spatial location for an object, state, or action.
5. Temporal Identification	Specifies temporal information about an action or state.
6. Degree Identification	Specifies the degree of an attribute, location, or time.
II. Event Specifying (Case) Relations	
1. Case Operations	Specify information missing from event frames or transform event frames into event frames of different types.
2. Event Generation	Specifies new events that incorporate given objects or propositions.
III. Relations Among Propositions	
1. Macrostructure Operations	Derive superordinate or subordinate propositions from given propositions.
2. Relative Operations	Specify metric and nonmetric relationships among objects and how states or events are related in space and time.
3. Dependency Operations	Establish causal, conditional, and logical dependencies among propositions.
4. Truth Value Operations	Modify the truth value of a proposition.

[a] Frederiksen et al. (1978).

Table 3
Frederiksen's Functional Contexts of Inference in Reading Comprehension

Types	Functions
First Stage Inference Resolution of Ambiguity Resolution of cataphora Dietic Inference: person, place, time	Interpretation of a current sentence by replacing anaphoric elements in proposition, with referrents and resolving ambiguities by selecting a preferred reading.
Connective Inferences	Connecting disconnected propositions.
Extensive Inference	Generating new propositions which extend meaning given by original set.
Structural Inference	
Segmentation Topical Inference Reduction	Segmenting and organizing a text, building a coherent model of a text as a whole.

[a] Based on Frederiksen, 1977b; Frederiksen et al., 1978.

Discussion

From the discussion of these frameworks, it becomes apparent that the diverse role of inference is still rather pervasive. Certainly the various frameworks approach a working definition of certain inference types and processing strategies; however, the categories and processing strategies within each framework seem to be more descriptive than prescriptive. As Frederiksen (1977b) suggested, a theory of inference in reading comprehension must not only "specify the types of inferential operations which occur and the discourse contexts in which they occur, but it also must account for the processes which control inference." Further, it must "account for the interplay of text-based and schema-based information in the control of discourse processing." In this regard, each author converges on a notion that seems to have some implicit, prescriptive power. Namely, within the definition of relevancy pursued by Trabasso, the functions suggested by Frederiksen, and the variable constraints discussed by schema theorists, is the implicit suggestion that researchers and theorists need to pinpoint the processes used in reading.

In many respects, this implicit pursuit is consistent with elaborated notions of some schema theorists. Neisser (1976) stated, "The schema is not only the plan but also the executor of the plan. It is a pattern of action as well as a pattern for action" (p. 56). Winograd (1977) suggests that readers possess a large and rather diverse set of schemata for planning their schema activity in accordance with the communicative situation. Clark (1977) argues that the nature of the writer's assumptions of the reader and the reader's assumptions of the writer will have a variable influence on inferencing across readers and reading situations.

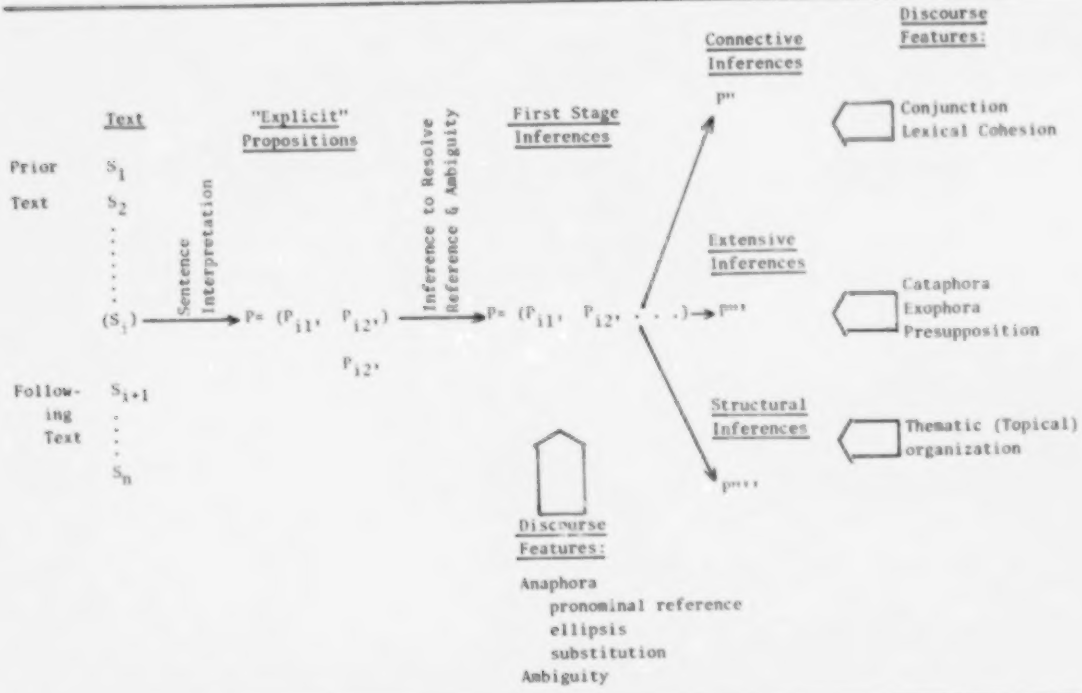

Text
"Explicit" Propositions
Inference to Resolve Reference & Ambiguity
First Stage Inferences
Connective Inferences
Discourse Features:
Prior Text
S_1
S_2
Sentence Interpretation
(S_i)
$P = (P_{i1}, P_{i2},)$
$P_{i2},$
$P = (P_{i1}, P_{i2}, \ldots)$
P''
Conjunction
Lexical Cohesion
Extensive Inferences
P'''
Cataphora
Exophora
Presupposition
Following Text
S_{i+1}
S_n
Structural Inferences
P''''
Thematic (Topical) organization
Discourse Features:
Anaphora
pronominal reference
ellipsis
substitution
Ambiguity

Figure 4. Frederiksen's levels of inference from discourse (Frederiksen et al., 1978).

That is to say, the author's use of conventions will likely have a variable influence on discourse processing depending on the communicative situation which includes the contract between reader and writer resulting from schema involvement.

In many respects, this implicit pursuit also is consistent with what sociolinguists and anthropologists have sought for some time, namely to address the pragmatics of discourse comprehension. As urged by Malinowski (1923) in defining the context of culture (and encouraged by Halliday with the terms "register"[2] or "context of situation"), "The more specifically we can characterize the context of situation, the more specifically we can predict the properties of the text in that situation" (p. 22). Or as vanDijk (1976) suggested:

> Both in linguistics and poetics it now appears, however, that many properties of sentences and discourses, and a fortiori of verbal communication in general, cannot be accounted for in terms of grammatical theories of the usual kind. What is needed is a pragmatic component in which rules, conditions and constraints can be formulated based on systematic properties of [speech acts] and communicative contents. (p. vii)

The diverse frameworks for inferences proposed by Collins et al. (1977), Trabasso and Nicholas (in press), Warren et al. (in press), and Frederiksen (1975, 1977a, 1977b; Frederiksen et al., 1978) provide evidence of the progress that has been made toward understanding the nature of reading comprehension. At the same time, their failure to seriously address a contextual feature is delimiting. Certainly it would seem that further analyses should incorporate a more thorough examination of the pragmatics of discourse processing. Indeed, an examination of pragmatics may lead to a model for models of inferencing—a model that will account for the interplay of text, reader, and context. The intent of this paper is not to describe such a model but to support efforts in this direction.

References

Anderson, R. C. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Chafe, W. L. *Meaning and the structure of language*. Chicago: University of Chicago Press, 1975.

Clark, H. H. Inferences in comprehension. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

Collins, A., Brown, J. S., & Larkin, K. M. *Inference in text understanding* (Tech. Rep. No. 40). Urbana, Ill.: University of Illinois, Center for the Study of Reading, 1977.

---

[2] "The register is a set of meanings, the configuration of semantic patterns, that are typically drawn upon under specified conditions, along with the words and structures that are used in the realization of these meanings." Halliday & Hasan (1976, p. 23).

Dijk. See vanDijk, 1976, 1977.

Fillmore, C. J. The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart, & Winston, 1968.

Fillmore, C. J. Types of lexical information. In D. Steinberg & L. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*. Cambridge: Cambridge University Press, 1971.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458.

Frederiksen, C. H. Discourse comprehension and early reading. In L. Resnick & P. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, N.J.: Erlbaum, 1977. (a)

Frederiksen, C. H. Inference and the structure of children's discourse. Paper for the Symposium on the Development of Discourse Processing Skills. Society for Research in Child Development, New Orleans, 1977. (b)

Frederiksen, C. H. Structure and process in discourse production and comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Erlbaum, 1977. (c)

Frederiksen, C. H., Frederiksen, J. D., Humphrey, F. M., & Ottesen, J. Discourse inference: Adapting to the inferential demands of school texts. Paper presented at the American Educational Research Association Annual Meeting, Toronto, 1978.

Halliday, M. A. K. *Language and social man*. London: Longman, 1974.

Halliday, M. A. K., & Hasan, R. *Cohesion in English*. London: Longman, 1976.

Kintsch, W. Comprehension and memory of text. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*, Vol. 5. Hillsdale, N.J.: Erlbaum, in press.

Leech, G. N. *Toward a semantic description of English*. Bloomington: Indiana University Press, 1969.

Malinowski, B. The problem of meaning in primitive languages. Supplement to C. K. Ogden & I. A. Richards, *The meaning of meaning*. London: Kegan Paul, 1923.

Neisser, V. *Cognition and reality*. San Francisco: Freeman, 1976.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Rescher, N. *Hypothetical Reasoning*. Amsterdam: North-Holland Publishing, 1964.

Schank, R. C. Conceptual dependency: A theory of natural language processing. *Cognitive psychology*, 1972-1973, 552-561.

Schank, R. C. Computer models of language: Identification of concepts underlying natural language processing. In R. C. Schank & K. M. Colby (Eds.), *Thought and language*. San Francisco: W. H. Freeman, 1973.

Simon, H., & Rescher, N. Cause and counterfactual. *Philosophy of science*, 1969, *33*, 323-340.

Simmons, R. F. Some semantic structures for representing English. In R. Freedle & J. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.

Spiro, R. J. Remembering information from text: Theoretical and empirical issues concerning the state of schema reconstruction hypothesis. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Trabasso, T., & Nicholas, D. W. Memory and inferences in the comprehension of narratives. In F. Wilkening & J. Becker (Eds.), *Information integration by children*. Hillsdale, N.J.: Erlbaum, in press.

Warren, W. H., Nicholas, D. W., & Trabasso, T. Event chains and inferences in understanding narratives. In R. Freedle (Ed.), *Discourse processing: Multidisciplinary approaches*. Hillsdale, N.J.: Erlbaum, in press.

Winograd, T. A framework for understanding discourse. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

vanDijk, T. A. *Pragmatics of language and literature.*
Amsterdam: North Holland Publishing Company, 1976.

vanDijk, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

Robert J. Tierney, Ph.D., is associate professor of education, University of Arizona, Tucson. Joseph L. Vaughan, Jr., Ed.D., is assistant professor of Education, University of Arizona, Tucson. Connie A. Bridge, Ph.D., is assistant professor of education, University of Kentucky, Lexington.

Reading is Recognition—When Reading is not Reasoning

J. Jaap Tuinman

> Meaning, as I think of the terms, is the result of the application of prior codings of the organism to the present decoding (i.e. reading) task. Wendell W. Weaver, 1977.

When someone reads a story these questions can be asked:

- What must the reader comprehend?
- What does the reader comprehend?
- What can the reader comprehend?

By and large efforts to conceptualize the measurement of comprehension have concentrated on the last one of these questions at the neglect of the first two. In the final section of this paper it will be argued that as a result we have mistakenly come to believe that "reading is reasoning," whereas, in fact, much comprehension activity can be best characterized as recognition.

Those persons concerned with the improvement of practices in the measurement of reading comprehension face two immediate challenges. First, criterion referenced measurement of comprehension skills is in a chaotic state. Tests proliferate in the absence of either theoretical justification or adequate psychometric undergirding. One cannot fault practitioners for their actions; neither can one void the need for such action. Nevertheless, the large array of "criterion-referenced" tests marketed and used rests on very shaky assumptions indeed. Second, cognitive psychology's rather recent preoccupation with both learning and reading has resulted in a number of insights that require assimilation in the conceptualization of the measurement of reading comprehension.

The first part of this paper will briefly elaborate on some of the problems with criterion reference measurement of reading comprehension skills. Some of the modifications of current thinking about measurement of comprehension suggested by recent formulations in cognitive psychology subsequently will be discussed.

Lumsden (1976) argues that educational measurement per se is at an impasse. When viewing its application in reading comprehension this author must agree, in spite of increasing mathematical sophistication in establishing indices of reliability, validity, etc. The problem, unfortunately, is not with the algebra but with the thinking behind it. This, of course, is not a new problem. Throughout the yet brief history of educational measurement influential writers such as Richardson (1936) and Buros (1948) have cautioned against the mistake of confusing the proliferation of mathematical superstructures with genuine progress in laying solid foundations.

The point, unfortunately, is well illustrated by current work on technical concerns in criterion referenced measurement (CRM). Illustrative of the problematic nature of much of the activity is a recent paper by Berk (1978). Specifying a 95% confidence interval, he is able to show that one needs 58 items to test the comprehension of a population of 1,000 sentences if one uses 80% mastery as the criterion score. This sounds very precise and helpful until one realizes that Berk's calculations are based on Bormuth's unproductive conceptualization of the measurement of comprehension (Mehrens, 1970; Tuinman, 1970).

That the development of new indices does not necessarily even reflect progress on the technical side is demonstrated by Downing and Mehrens (1978). When these authors compared six single-administration reliability coefficients for CRM's, they found that only one measured a test characteristic that differed from the classical Kuder-Richardson formula's.[1] More disturbingly, when Smith (1978) compared five popular item selection methods he found that none of the methods was consistently superior. Indeed, random selection of items worked just as well as any other method.

The general point that the elaboration of mathematical techniques can obscure basic conceptual problems is yet more clearly demonstrated when one reviews recent work on establishing adequate criterion, mastery, or pass scores. The intent, of course, is to formulate techniques that will validly classify individuals as masters or non-masters. Two of the more interesting papers by Huynh (1977) and Faggen (1978) illustrate current attempts to deal with the issue and inadequacy of the techniques.

The procedures for establishing mastery scores are elegant and appear effective. Nevertheless, both papers share a very fundamental problem. Whether or not a student is validly classified as a master on test *i* is judged by his or her performance on test *j*. Presumably, test *j* is more complex, more general or higher in some skills hierarchy than test *i*. The author grants that adequate tests of accuracy of classification are needed. The bulk of the unresolved issues, however, relate precisely to the specification of the relationship between tests *i* and *j* and to the adequacy of test *j* as a criterion.

This is far from being an issue for idle contemplation. It raises its head in every reading management system currently in use. As test constructors we are saved by the fact that most tests, whatever their label, format, or content, load heavily on a general verbal factor. Were it not for that fact, teachers would be confronted with many more puzzling patterns of performance on series of tests in their "management" batteries than they already face.

---

[1] This study did not include Brennan and Kane's (1977) index. However, see Lumsden (1976) for a critical evaluation of the signal/noise ratio indices.

Seldom are relationships among tests made explicit. Implicitly, however, hierarchical connections and transfer of acquired skills to other more complex ones are routinely assumed. Criterion-referenced comprehension tests fail, however, precisely because either they cannot satisfy these assumptions or, in the event they do, there are precious few ways to demonstrate the fact empirically.

An early optimism about the possibility of isolating and describing learning hierarchies is now tempered by the realization that even in the case of simple behaviors, empirical verification of hierarchical relationships is difficult. Two recent papers demonstrate the paucity of effective techniques. Guay and McCabe (1978) point out serious shortcomings in existing tests of hierarchical dependency. They then present a test which, they claim, remedies these problems. However, being restricted to a pair-wise comparison of skills, their test, too, is very limited in scope. The state of the art is best illustrated by Griffiths and Cornish (1978). They extended White and Clark's Test of Inclusion to more than three items per skill. The relationships among seven skills, representing the development of an introductory chemistry concept, were then analyzed. In all, some 10 possible hierarchies result from the analysis—hardly a comforting thought, considering the relatively simple nature of the skills involved. The prospect of extracting ourselves from the measurement muddle through specifying skills hierarchies in the comprehension domain is dim indeed.

Even if more powerful statistical techniques were available and practical, there is still grave doubt regarding the applicability of the hierarchy concept to the comprehension domain. Every teacher knows that it is possible to ask a very easy "higher order" question and a very difficult "lower order" question. If skills hierarchies are to be studied at all, this should be done within a given level of language, within a defined lexicon. We will never understand the contribution of, for instance, "being able to use context" to "being able to draw inferences" unless we first systematically study the relationships among such skills for a particular set of lexical elements, syntactic patterns, semantic domains.

Part II

It is perhaps important to admit that "skill" may be an inappropriate metaphor for much of what readers do with language. In the distinction between "skill in" and "knowledge of" language, reading educators have tended to emphasize the former. It is possible, however, to view "drawing an inference" as "knowing an inference" under certain conditions. Pace (1978) studied children's comprehension of stories with familiar and unfamiliar content. In discussing her results she makes the following comment:

> Even kindergarten students can answer questions about the implicit forms of familiar stories. Thus, they *look like* they are making inferences because they go beyond the information given. However, since these same kindergartens apparently do not operate on the explicitly provided information in stories

> about unfamiliar situations to enhance their understanding of them, they are not likely to be using deliberate inferential processes. Once they recognize that a story concerns something that they know, they may simply use their own knowledge—their own "scripts"—to supply effortlessly information that was omitted from the story itself. (p. 12)

Pace provides a clear example of a reading behavior that looks like reasoning but is in fact an instance of recognition.

In 1917 Thorndike made his well-known claim that reading is reasoning. At the same time he initiated a measurement tradition which made his claim nearly unfalsifiable. Had Thorndike been a rethoritician concerned with structure of narratives he might have made the leap and asked how structure variables would relate to comprehension. But he was not and those questions did not arise in full force till some 60 years later. Instead of stories, Thorndike chose brief paragraphs to study and measure comprehension. This, in the author's opinion, has had some most unfortunate consequences for the teaching and testing of reading comprehension.

Thorndike set a measurement tradition that has saddled us with a vexing conflict: many test items ask questions, often very clever ones, about story information that tends to be totally neglected, overlooked as of no interest when reading anything except reading tests. There are pragmatic reasons why reading tests exploit trivial information. Tests must be limited in bulk for "obvious" reasons. Therefore, most tests use a large number of exceedingly brief paragraphs or a lesser number of somewhat larger passages. In either case, to achieve respectable reliability, relatively large numbers of items are written. The combination of many questions asked of limited textual material frequently leads to a near exhausting of the information available.

There is a second practical consideration for test writers in this regard. Often items must be of a specific kind, such as "inference questions." Many passages provide very restricted opportunities to ask such questions. Hence, item writers frequently are forced to pounce on trivial information if it provides an opening for a question requiring inferencing. The problem is not only that unimportant information forms the basis for questions but also that so much of the content of the passage is used in this fashion. The text is milked dry; the reader is forced to process it in a manner atypical of much natural reading.

One can conceivably argue that such tests are useful to reveal potential limits of a reader's comprehension even though under many reading conditions he may not pay attention to most of the details tested either for the purpose of remembering them or as take-off points for elaborations.

Notice that this argument acknowledges that these tests, then, emphasize what the reader *can* comprehend at the price of ignoring what he *does* comprehend. The argument, moreover, has merit only as long as one acknowledges that the shift from actual to potential comprehension may, and the author will argue does, entail a qualitative difference in the product of the comprehending.

Flavell (1976) crystallizes much of the perspective of current cognitive psychology on intellectual functions when he remarks:

> In any kind of cognitive transaction with the human or nonhuman environment, a variety of information processing activities may go on. Metacognition refers, among other things, to the active monitoring and consequent regulation and orchestration of these processes in relation to the cognitive objects or data on which they bear, usually in the service of some concrete goal or objective. (p. 232)

The notion that the reader is an active processor of text, monitoring the information encountered, interactively assigning and reassigning meaning is central to the cognitive psychologist's view of reading. An important question, therefore, is how effectively a particular reader regulates his understanding of what is being read; how sensitive is the reader to multiple interpretations; how efficient is he in selecting appropriate ones; how attuned is he to relative loss of meaningfulness of the larger message when inappropriate meanings are assigned to its components. When, in tests, we ask questions of readers we replace internal monitoring with an external force. The fact that the reader when prompted externally *can* find a fitting answer, *can* detect a semantic incongruity, *can* be made aware of a contradiction, tells us little about what happens to the reader's comprehension when the only prompt is internal.

Cognitive psychologists have added the terms "intention" and "goal" to the lexicon of motivation and purpose. Intentions and goal shape the interaction between reader and text, codetermine level and speed of processing.

Intention and goal are part of and arise from the larger pragmatic context of the communicative act. Too often, the author believes, has measurement of comprehension been conceptualized without reference to this context.

To elaborate this claim, let me introduce three, somewhat arbitrary, contexts for comprehending reading.

1. Private comprehension: the reading of materials without further demands on the information contained in the text.

2. Communicative comprehension: the reading of materials containing information likely to be informally shared with others.

3. Formal comprehension: the reading of materials containing information to be formally shared with others.

Private comprehension. A few moments before starting the writing of this section of the paper, the author finished the hurried reading of yet another best seller, Gerald Seymour's *Kingfisher*. The narrative is fast, packed with action both physical and psychological. The story has strong forward pull; it is difficult to put down and, under time constraints, invites a fair amount of skipping marginal paragraphs. Guilt, rising from giving priority to Seymour's writing over my own writing added to the speed of my reading the novel. I calculated that I read the last 150 pages at over 600 wpm.

If anyone were to query me as to specific events or existents (e.g., gestures, clothes, thoughts), not only might I not know appropriate answers, but I also know that I did not attend to many of those details. I did not find them important. They were not relevant to the kind of comprehension I required for enjoying the novel. Moreover, I read the book not expecting to have to share that kind of information with others. I did not "milk" the story, did not have a pragmatic (in a discourse sense) or esthetic (as in the case of reading Shakespeare) reason to do so. The essential characteristic of private comprehension is that it both results from and satisfies inner information needs.

It must be emphasized that the focus of my argument is not on the distinction between product and process of comprehension. The point is that typical tests of the product of comprehending zero-in on products that are atypical rather than typical of certain types of reading frequently engaged in.

A by-product of Thorndike type measurement, however, is indeed the neglect of the study of comprehension processes in naturalistic settings. How do readers, for instance, differ in their processing of story information; how do they differ in their selection of salient information; in their need to understand; in their tolerance for ambiguity, redundance, etc.? How do readers grow as private comprehenders?

Communicative comprehension. A good example of what I have in mind here is the reading of an account of a sports event. There is an unspoken expectation that I may share some of the information. The probability that the event will come up in exchanges with fellow enthusiasts is high. Moreover, I may be already partly informed. I heard a newscast; I read a "preview of the event" type analysis; I discussed the upcoming event with friends.

I suspect that reading in this context differs from private reading in at least two dimensions: (a) the amount and detail of pre-reading information available, and (b) the expectation that the information read will be shared in communication.

Note that I use the term "pre-reading information," not script, frame, schema, etc. Private comprehension, too, presupposes scripts. The kind of pre-reading information involved in communicative comprehension is specific, factual knowledge rather than general, abstracted, structural knowledge.

If my distinction has merit, materials that ordinarily are read in the pragmatic context of informal communication tend to be more known. As a consequence we may expect a different kind of processing than takes place in either private or formal comprehension.

Though certain types of writings may be more frequently associated with one type of context, the distinction is far from absolute. When working with colleagues who are all avid readers of "best sellers," I may develop an informal communicative comprehension set when reading novels. I may discuss characters and events before reading about them. I may expect to contribute to staff room talk about details of the story.

The expectation that information read for communicative comprehension will be shared at a later date is not formal, well-defined, deliberate. Yet, who has not come across an odd bit of information about a well known public figure and decided to remember it *to tell later?* "Did you know that Carter. . . .," for example. This behavior, this processing strategy, does not tend to exist during private comprehension—"Did you know that Elg Schwartz (the hero in Schweig's best seller *Come in the Rain)* dyed his hair?" is just not a statement readers are likely to make.

Formal comprehension. This category is probably most suited to present practices of testing. The reading of a social studies chapter in a textbook is a good example. Here the reader's orientation and the external prompts in the form of questions are probably not too far apart. The test duplicates the process the reader has already gone through. The reader is aware that he has to "understand" every detail and "relate" it to the other information presented. Specifics are important; generalizations must be recognized; summaries are prepared physically (or merely in the mind) to be presented later. Fine discriminations among facts, events, motives, abstractions may be crucial.

One of the most curious developments in the measurement of reading comprehension is that concepts and techniques suitable for the measurement of formal comprehensions have come to be used for the measurement of all comprehension. The typical set of test items presented below demonstrate that contention unequivocally.

> Janet and Jim went for a short walk. They went for about ten minutes. Janet wore boots, but Jim was wearing sneakers. Mother did not go along. She was cooking dinner.
>
> 1. The children went for a short: ride walk flight
> 2. Janet wore: a skirt sneakers boots
> 3. Who did not go for a walk? Mother Janet Jim
> 4. The children couldn't stay out because they had to: go to bed clean up their room eat

To repeat, the issue is not whether these questions measure what a young reader can understand. Rather, the concerns with a measurement tradition exemplified by these kinds of items can be summarized as follows:

- neglect of what the reader does comprehend as he reads in various types of contexts versus what he can be made to understand when faced with external prompts in the form of questions.

- neglect of research on techniques for measurement of typical products in a variety of natural reading contexts.

- as a corollary, neglect of the study of the psychological processes associated with obtaining such typical comprehension products and of the instructional interventions to further their attainment.

Part III

Referred to above is the fact that Thorndike's claim that reading is reasoning was made nearly unfalsifiable in the context of the measurement tradition which developed in the years following his pronouncement. Elsewhere (Tuinman, 1972) this author has argued that Thorndike's view of reasoning, typically associationistic, probably was narrower than ours. Nevertheless, once test constructors included inferential comprehension items in test batteries it naturally became increasingly impossible by any technique of statistical reduction to show anything but that reading was indeed reasoning. The work of Davis (1944, 1968, 1972) nicely illustrates the premises, the procedures, and the unavoidable outcomes. This point is well known and well argued. There are two further facets to the issue, however.

One direct consequence of the short paragraph-many questions paradigm coupled with the necessity for many levels of questions (in a taxonomic sense) is the concentration on ever finer discriminations. Anyone who has written comprehension test items knows the tortuous process of "having to find one more angle," so one more question may be asked of a specific paragraph or passage. Conceptual distinctions are made for the sake of asking a question, and perspectives on the information contained are pursued that would not normally occur to any reader as a reader (as opposed to tester). The question, therefore, is whether reading is reasoning because it is reasoning or because we invest so much of our reasoning as testers into its measurement.

Neisser (1976) develops a somewhat analogous argument for the case of intelligence. He holds that psychometric intelligence tests basically measure academic intelligence. It is a mistake, he says, to identify academic intelligence with intelligence *per se*. Test constructors, by definition, are people who were good students, academically successful. As Neisser puts it: "tests are devices for finding other people who resemble their creators."

To what degree is reading reasoning because of the clever minds of those who construct reading comprehension tests? To what extent have decades of cross-validation of tests also been decades of conceptual in-breeding; decades of an unconscious process of boosting the reasoning factor and minimizing other aspects of the task?

Thorndike's first reading test questions have long been the prime vehicle for measuring comprehensions (as opposed to, e.g., summaries, paraphrases, retellings), and this practice has naturally led to the use of multiple choice items in comprehension tests. This convention, I believe, has seriously exacerbated the "imposed reasoning" problem. Multiple choice tests require response alternatives that are credible and yet not true or correct. Consequently each alternative shares certain conceptual features but not others with the correct response. The process of deciding among alternatives has a name: reasoning. Once more the measure used colors the intellectual operation of interest. To what extent is reading reasoning because of the reasoning processes built into the creation of clever distractors? Traub and Fisher (1977) analyzed the equivalence of constructed-response and multiple-choice tests across two sets of mathematical reasoning and two sets of verbal comprehension items. The specifics of their study, aiming to identify format differences—other than those due to different errors of measurement, units of measurement, and arbitrary origins of measurement—are not particularly relevant to the purpose of this discussion. Their conclusion, however, is important.

> The tests of mathematical reasoning that were employed were equivalent regardless of format, but the tests of verbal comprehension were not. In particular the free-response test of verbal comprehension seemed to measure something different than standard multiple-choice and Coombs test of this ability. . . (p. 367)

It would be interesting, indeed, to apply the methodology employed by Traub and Fisher to current comprehension tests to ascertain whether such format differences could be replicated and whether they would be interpretable in terms of an "added-on" reasoning factor.

Comprehending reading is (a) a recognition of propositions stating a fact about the world—including facts that are statements of relations among facts, (b) a "reasoning out," an inferring of such facts, or (c) a mixture. In most cases of reading point *c* holds. Facts about the world can be either known or inferred from what is known. If reading is the mechanism to call such facts to the reader's attention it is a process of recognition in the former case and a process of reasoning in the latter. All conceivable inferences can be either inferred, if they are not "known," or recognized, if they are.

A view of reading comprehension that invokes schemata, scripts, or frames, by necessity must curb the role of reasoning in certain contexts. Consider the conclusions that Rumelhart and Ortony (1977) draw about the relationship between cognitive schema theory and comprehension:

> Providing information in a structured form most closely resembling the structure of the schema which will be required for its interpretation maximizes the likelihood that the interpretation will be appropriate and *minimizes the processing required.* (italics added; p. 131)

In other words, under conditions in which the schema is very complete or in which the text contains little new information, processing is minimal. Under those conditions reading becomes less reasoning and more recognition. The reverse is true when the gulf between existing schemata and information presented in and implied by a text widens. Then reading becomes less recognition and more reasoning.

Pace (1978) has concretized this view in the study referred to above. Consider further the following illustration to reinforce the point.

> John is 6'6", a giant basketball player. Mary, standing 5 feet watches him in awe. So does Ken, who knows that Mary never again will look up to him as she used to.

The reader of this paragraph can infer the fact that Ken is taller than Mary either from general knowledge (i.e., the probability of men named Ken to be less than 5' is small) or from the linguistic cue, "never again will look up to him as she used to." Reading is reasoning!

Or is it? Imagine that the narrator stands in a television studio with John, Mary, and Ken. They all look at the story, projected on the wall in an effort to do some editing for a human interest public broadcasting story. What if someone would ask how tall Ken is? Would any reasoning or inferring take place?

The example is extreme. Rarely does a narrative mirror a known reality so closely. Yet, when a reader reads about realities with which he is familiar he can either infer unstated aspects of that reality from the text, or, with far less effort, supply them from his memory. Then, reading is not reasoning and neither is there any reason it should be.

This author has argued that current mechanical problems in the measurement of comprehension cannot be resolved until we reexamine our view of comprehension as skill, rather than knowledge. The author has tried to elaborate the view that our perspective on comprehension has been narrow, unduly circumscribed by our preoccupation with the assessment of what can be comprehended using the vehicle of information-exhausting questions on passages void of genuine content. Those arbitrary contexts for comprehending were proposed in an effort to illustrate that different information processing needs direct different processing strategies and conceivably result in different products of comprehension, unaccessible through the use of questions as external prompts. It was argued that reading has come to be viewed as reasoning in part because we have inadequately separated the product measured and the process of measuring. Finally, it was contended that a more appropriate view of reading characterizes comprehension as a process that involves both recognition and inferences about facts stated and unstated. Reading requires reasoning when either the linguistic code is complex or when the reader's schemata are inadequate to accommodate the text's information structure. When the message is coded in known linguistic structures (phonetic, syntactical, or semantic), and the text's information structure matches the reader's schemata, reading is merely recognition.

References

Berk, R. A. *Item sampling from finite domains of written discourse.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Brennan, R. L., & Kane, M. T. Signal/noise ratios for domain-referenced tests. *Psychometrika,* 1977, *42*(4), 606-625; 1978, *43*(2), 289.

Buros, O. K. *Criticisms of commonly used methods of validating achievement test items.* Proceedings of the 1948 Invitational Conference on Testing Problems. ETG., 1948, 18-20.

Davis, F. B. Fundamental factors of comprehension in reading. *Psychometrika,* 1944, *9,* 195-197.

Davis, F. B. Research on comprehension in reading. *Reading Research Quarterly,* 1968, *3,* 499-545.

Davis, F. B. *The literature on research in reading with emphasis on models.* New Brunswick, N.J.: Rutgers University Press, 1972.

Downing, S. M., & Mehrens, W. A. *Six-single-administration reliability coefficients for criterion. Reference tests: A comparative study.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Faggen, J. *Decision reliability and classification validity for decision-oriented criterion referenced tests.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Flavell, J. H. Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence.* Hillsdale, N.J.: Erlbaum, 1976.

Griffiths, A. K., & Cornish, A. G. *An analysis of three recent methods for the identification and validation of learning hierarchies.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Guay, R. B., & McCabe, G. P. *A chi-square test for hierarchical dependency.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Huynh, H. Two simple classes of mastery scores based on the beta-binomial model. *Psychometrika,* 1977, *42*(4), 601-608.

Lumsden, D. Test theory. *Annual Review of Psychology,* 1976, 223-258.

Mehrens, W. A. Scientific test construction—Pure and sterile. Review in *Contemporary Psychology,* 1970, *15*(11), 666-667.

Neisser, U. *Cognition and reality.* San Francisco: Freeman, 1976.

Pace, A. J. *The influence of world knowledge on children's comprehension of short narrative passages.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Richardson, M. W. Notes on the rationale of item analysis. *Psychometric,* 1936, *1,* 69-78.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. Anderson, R. Spiro, & W. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Erlbaum, 1977.

Smith, D. U. *The effects of various item selection methods on the classification accuracy and classification consistency of criterion-referenced instruments.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Traub, R. E., & Fischer, C. W. On the equivalence of constructed response and multiple-choice tests. *Applied Psychological Measurement,* 1977, *1*(3), 355-369.

Tuinman, J. J. *Bormuth's views on the theory of achievement items–Some questioning comments.* Paper presented at the annual meeting of the National Reading Conference, St. Petersburg, December 1970.

Tuinman, J. J. Determining the passage dependency of comprehension University, Vancouver, British Columbia.

J. Jaap Tuinman, Ph.D., is professor of education, Simon Frazier University, Vancouver, British Columbia.

Overt Indications of Reading Comprehension: Product and Process Considerations

William D. Page and Richard T. Vacca

Reading educators generally agree that reading comprehension is the major purpose of reading instruction; but, when it comes to defining reading comprehension, no single definition compels agreement from every corner. Reading comprehension is related to meaning. One approach to reading comprehension suggests that it is getting an author's meaning from the author's printed language. However, definitions of meaning that elicit agreement are as difficult to come by as definitions of reading comprehension. Smith (1971) observes, "The attempt to find a single definition for all the various idiosyncratic aspects of meaning is a rock on which many theoretical vessels have foundered" (p. 185). Whether or not theoretical vessels founder, reading instruction continues to plunge through educational waters with its ultimate course set for comprehension.

The lack of a single compelling definition of reading comprehension that serves all concerned, whether they be readers, teachers, or test makers, is reason enough for researchers to study it. Disciplines begin and proceed because of unanswered questions. The apparent characteristics of reading comprehension are sources of questions. Reading comprehension is internal, it goes on inside the organism. As such, it is not directly observable, and therefore, not directly measurable. Its existence can be inferred from introspection and from a framework of overt indicators. Reading comprehension is a dynamic process, not an object. It is more an activity of thought than a possession. Most tests suggest that reading comprehension is one subskill of reading among many. However, if an individual reconstructs an author's message from printed language, we are satisfied that reading is occurring, regardless of performance on other assumed subskills of reading. Conversely, no matter how well an individual performs on tests of assumed subskills of reading other than comprehending printed language, if that individual can not understand the author's message we are hard put to say that the individual is reading. Reading comprehension is not a subskill of reading, it *is* reading.

The usual procedure for interpreting a battery of tests of skills is to focus on the composite score. Obviously, this can be misleading. Composite scores tend to be averages or weighted combinations. The characteristics of reading comprehension require a theoretical framework other than merely averaging a collection of tests of skills if we are to make sense of the covert process we seek to understand.

Pearson and Johnson (1978) create a medical simile that generates some insights:

> Examining reading comprehension is much like trying to diagnose a respiratory disease. A doctor will perform a blood test. . . . Then, he or she might take an X-ray and attempt to match the patient's X-ray against standards that typically indicate a certain disease. . . . Throat cultures, body temperature measurements and other vital signs are assessed in order to gather more and more indexes to assist in the diagnosis. . . . Each test is a new perspective, not completely independent from the other tests, but adding new information nevertheless. (pp. 154-155)

The medical doctor can not rely on any single indicator of the internal state of affairs in the organism; rather, he or she constructs a corroborative framework of interrelationships among the various indicators of the patient's condition. Similarly, we have a variety of indicators of reading comprehension available.

Making sense of the indicators of reading comprehension can be compared to reading and interpreting a series of figures that provide context, a frame, for one another. Each reading on a gauge gives us a clue but we must understand what the gauge is doing, how it relates to the indications on other gauges, and what it is we are seeking in order to render a useful interpretation. We must avoid confusing the gauge with the phenomenon we seek to understand. Further, we must avoid the error of working on the gauge rather than the phenomenon to create the situation we desire. The needle of the fuel gauge of an automobile tells us to some degree about the internal state of the fuel tank. The gauge tells how much fuel the tank contains. When the fuel gauge needle points to empty, we assume that the car will stop unless we add fuel. We can break the protective glass on the fuel gauge and bend the needle over to full, but that will not change the state of affairs in the fuel tank. The car will still stop when it has run out of fuel. Similarly, no medical doctor would tamper with a blood sample to produce the desired outcome of curing the patient. If the fuel gauge remains intact, we can make a useful inference about the internal state of the fuel tank. Indicators of reading comprehension function similarly. We get information about the internal state of the organism that permits us to infer, with a good deal of error, to what degree the individual is comprehending the author's printed language.

The error involved in inferring to what degree comprehension is occurring in the organism is a major concern. Carroll (1972) suggests, ". . . most presently available tests do not permit a satisfactory assessment of an individual's 'absolute' level of comprehension ability" (p. 3). Tuinman (1973) demonstrates that subjects who have not read the paragraphs to which questions refer in five major, widely used standardized tests answer the questions about as well as subjects who have read the paragraphs. Bormuth (1970) points out that test makers seldom adequately explain the rationale underlying the processes used to write and select comprehension questions in tests. Tyler (1974) indicates that reading comprehension tests probably do not test exactly what they say they test. It is sufficient to say that tests of reading comprehension—or rather, tests *labeled* "reading comprehension"—provide indicators that include substantial error.

The error involved in struggling to assess reading comprehension prompts several insights. Today, with school districts moving toward using mandated reading levels as ways of determining the promotion and graduation of youngsters, educators have a great need for accurate methods of assessing an individual's reading comprehension ability (Harper & Kilarr, 1978). The possibility of confusing the strength of the need to assess reading comprehension with the strength of our present technology for assessment is enormous. Our present technology is probably not good enough to meet the need for assessment of reading comprehension evoked by policies that mandate reading ability levels as determinants of promotion and graduation. The relationship between need to assess and capability for assessment sharpens the focus on the need for a corroborative framework of indicators of reading comprehension. We do not suggest that a corroborative framework will move technology to the point of satisfying the need for assessment. However, it provides a way of narrowing the gap between what consumers desire and what testing can deliver.

A corroborative framework of indicators of reading comprehension is required to make sense of an individual's performance in order to infer the degree to which the individual is able to reconstruct an author's message. Farr (1969) pinpoints a concern in assessment when he suggests that no single test should be used as the sole indicator of an individual's reading. The need for a corroborative framework of reading comprehension indicators is clear.

Knowledge Stating Within a Corroborative Framework

Reading comprehension often is measured in the same brutish manner as the sideshow bell-ringing contest. The advent of statewide proficiency and minimal competency testing alluded to previously has even managed to add a carnival atmosphere to this particular measurement attraction—with parents, legislators and educators watching from the sidelines to see how high youngsters bang the "basic skills disc" up the "literacy column."

Reading comprehension traditionally has been assessed through performance outcomes which represent overt indicators for a covert process. These outcomes signal, more often than not, product indications of comprehension. Product indicators emphasize how much the reader has learned as a result of comprehending passage content. The degree to which the propositional meaning of the passage has been demonstrated indicates the degree to which one has comprehended. The more one states knowledge about what has been read, the greater the indication of "strength" as a comprehender. Hence, the analogy to the bell ringing contest at the carnival!

The reading comprehender, as a result, is put in a situation where he or she is required to state knowledge through vocal or written speech, usually in the form of sentences. Gagne and White (1978) define knowledge *stating* as learning outcomes that rely on "the assertion of sentences in a form that preserves the propositional meaning of sentence inputs, including those of organized prose" (p. 189).

There are two aspects of knowledge stating which provide us with product indicators of comprehension: knowledge retention and knowledge transfer. Knowledge retention requires the reader's comprehending performance to be judged in terms of completeness and correctness, insofar as performance preserves the propositional meaning explicit in the passage. Knowledge transfer, on the other hand, requires the reader to make inferences about the prose material or to relate information in one passage to the information in another passage.

Questions are widely used devices to prompt knowledge stating. There are inherent limitations in the use of questions as overt indicators of reading comprehension, despite the fact that their use has been notable in the major prose learning research of the past two decades (Ausubel, 1968; Frase, 1975; Rothkopf, 1971).

Free recall, on the other hand, offers straightforward measures of comprehension. In the past several years free recall has proved immensely valuable as an indicator of comprehension. An analysis of recalls goes beyond product considerations. As Marshall and Glock (1978-79) indicate, "By studying recall for information presented in a passage, one is not just studying the process of comprehension for written material. One is also studying the process of thinking, or at least of that part of thinking that seems to be linguistically encoded" (p. 13). Recalls apparently provide insights into process that questions thus far, apparently, have not.

Some Limitations of Questions

Product indicators of reading comprehension attempt to assess performance with prose material through the answering of questions related to passage content. The products of comprehension typically are emphasized when questions are used to assess gains in information. No wonder, then, that Rothkopf (1972) characterizes the text learner as an "information eater." He has proposed that two empirical questions direct adjunct questioning research on prose learning: "How is daily gain in weight affected by what you place in the feed bin and by how often you do it?" (p. 331). These questions are posed in the language of the "information eater" metaphor. Page and Vacca (1978) suggest that performance indicators that look solely at "weight gain" are limited in terms of what they can tell us about the reading comprehender.

How one goes about the process of comprehension matters. The type of question used to indicate performance, for example, must be considered carefully. Rote learning of text, triggered by the use of verbatim questions, may merely represent a meaningless string of speech sounds (Anderson, 1972). "The trick," according to Anderson, "will be to devise techniques for constructing questions that can be answered if a person has semantically encoded a communication but not answered if it has been encoded perceptually or phonologically" (p. 148). Most of the written questions devised in the adjunct question research required verbatim responses. Readers were assessed on their accuracy to repeat literal bits of information. As Anderson and Biddle (1975) assert:

> Of practical interest is the part that adjunct questions can do more than increase the accuracy with which people are able to repeat strings of words. . . adjunct questions which entail paraphrase and application of principles and concepts to new situations may be especially facilitative. . . . (p. 103)

Knowledge transfer, we believe, has not received sufficient attention by researchers who contemplate the use of questions as indicators of comprehension performance.

Another limitation of questions is that they occur after the fact; questions test post reading performance. Carroll (1972) noted that post reading tasks confound comprehension with the effects of memory. Readers may indeed understand an author's message during reading, but will not preserve it over a period of time. A set of questions prompts knowledge stating, for sure, but the questions are cognitive triggers at best. They may or may not activate retrieval for some students. The appropriateness of questions to elicit a thoughtful response probably varies with the comprehender. In addition, readers who comprehend a communication during reading may be constrained by their facility to remember during post reading tasks. It is possible, therefore, for some readers to come up with a low score on a comprehension test even though they understand the passage during reading.

Memory, nevertheless, is inextricably linked to knowledge stating. Gagne and White (1978) define memory structure as the organized contents of long-term memory that result from learning. To the extent that information is stored in long-term memory in the form of propositions (as opposed to images or episodes), questions should attempt to indicate not only "how much is learned" but also "what is learned." All too often the use of questions as a gauge for comprehension performance underlies a one-stage model of information processing.

Mayer (1975) suggests that in a one-stage model, the critical learning condition focuses on how much of the presented material is received by the learner. A two-stage model of processing, on the other hand, assesses how much is learned within the context of prior knowledge already in the learner's memory structure. A three-stage model accounts for information processing not only in terms of how much is learned within the framework of prior knowledge, but also in terms of what aspects of the learner's existing knowledge are activated during learning and used as an assimilative set.

One-stage models of processing usually assess comprehension in a purely quantitative fashion. Such models tend to deny the existence of prior memory structure and also what aspects of the learner's existing knowledge are activated during learning and used as an assimilative set.

In short, there is a need to know more precisely how learning from text is structured in long-term memory—questions as measurement devices have not done this. Rickards' recent work (1976, 1977-78), however, points to new directions in the use of questions as overt indicators of comprehension performance. In addition, continued applications of discourse analysis (Frederiksen, 1975) which have spawned notable attempts to relate the structure of connected discourse to the structure of the reader's recalls (Marshall & Glock, 1978-79) should help to close the schism, to some degree, between our knowledge of the products of comprehension and its underlying processes.

In-Process Indicators of Comprehension

As noted, most of the tests of reading comprehension used in schools today are post reading tests, situations in which the reader's responses are elicited following reading. As such, these tests confound the effects of memory with reading comprehension. Carroll (1972) notes other factors that may interfere with indications of sheer language comprehension as deduction, inference, and problem solving. One clear division that is researchable is the distinction between responses elicited following reading (product indicators) and responses elicited during reading (in-process indicators). The latter, in-process indicators, promises fertile ground for research questions and the possibility of the development of feasible testing techniques for use in schools. There are many research approaches to reading comprehension as it occurs. Here, we focus on two approaches because they are presently relatively well developed, and promise to spawn further insights and applications in the near future. The cloze procedure is one approach and miscue analysis is a second. Others, not included here, are on the horizon and may be just as promising, but the two we have selected serve as examples for examining linkages between product and in-process indicators of comprehension.

Cloze Techniques

The history of the cloze performance in relation to readability is well documented back to Taylor's (1953) work. Bormuth (1969) shows that significant amounts of variance are shared between cloze scores and scores of multiple-choice question tests of literal comprehension. Bormuth (1975) provides grade-level equivalents of cloze scores for grades 3 through 12. The cloze form Bormuth (1969, 1975) uses is the conventional fifth-word deletion technique with only responses exactly corresponding to the deleted word scored correct.

Conventional cloze responses are elicited during an initial silent reading. As such, they are in-process indicators of comprehension. The concurrent validity of conventional cloze scores as indicators of literal reading comprehension is clear. However, several limitations are of concern.

The construct validity of cloze performance is in doubt. Horton (1973) shows that cloze scores share significant variance with indicators of subjects' performance on tasks of relating ideas and words and claims to have demonstrated the construct validity of cloze performance. Despite Horton's findings, the task of filling in blanks lacks some important characteristics of ordinary reading. When a critically important word that is not repeated in the text is deleted, the reader does not encounter an opportunity to fully reconstruct the author's ideas. The cloze task involves a clerical function, filling in blanks one after another, an aspect of performance that probably interferes considerably with the psychological processing by the reader.

Cloze performance shares variance with common intelligence test scores, and correlates with reading test scores about as well as intelligence test scores do. As a single indicator of reading comprehension, conventional cloze scores fall short of what is required to determine the academic destiny of a youngster. Despite the limitations of cloze scores, the cloze procedure is a useful tool for estimating readers' comprehension when inexpensive estimates of group performance are sought. With individuals, cloze performance can be used as one indicator among many in a corroborative framework of comprehension indicators. Few techniques rival the reliability and ease with which cloze tests can be constructed, administered, and scored for large groups using passages selected by a teacher or researcher. Conventional procedures for creating teacher-made tests of selected passages tend to be expensive and of questionable reliability (Bormuth, 1975).

With some of the limitations and advantages of cloze procedure clearly in mind, some recent developments with the cloze procedure warrant examination in relation to the notion of a corroborative framework of indicators of comprehension. Page (1975) demonstrates that scores of conventional cloze tests share significant variance with scores of post oral-reading cloze tests. A conventional cloze test is administered without prior contact of the subject with the passage, while a post oral-reading cloze test is administered immediately following an oral reading of the same passage by the same subject. Ganier (1976) finds similar results with a different sample and design. Carey (1978) shows significant shared variance among post oral-reading cloze test scores and standardized test scores. The post oral-reading cloze test is a useful indicator of comprehension that can be used following an oral reading, but the limitations of conventional cloze scores remain in effect.

Another aspect of cloze performance, whether it be with a conventional cloze test or a post oral-reading cloze test, is of concern in research design. The way a reader processes language differs depending on the size of the unit of language the reader elects to consider during reading. Readers may focus on bits of language such as phonemes, graphemes, morphemes, and words. Readers may focus on phrases, clauses, and sentences; or they may focus on paragraphs, stories or documents, whole books, or on relationships among universes of discourse as in metalinguistics. The characteristics of the cloze task might tend to keep the reader's score of operations focused on words and phrases, and on the semantic and syntactic aspects of units of that size, rather than on the smaller or larger units. Clearly, the phonemes and graphemes of deleted words in a cloze test are not available to the reader, thus ruling out cloze as a strong indicator of performance with bits of language. The lack of information due to deleted key words takes its toll on the reader's opportunity to construct an accurate web of interrelated propositions representing the passage being processed. As a result, cloze performance probably taps the middle sized units of language. We speculate here that cloze performance is probably a poor choice as a dependent variable in research seeking to assess the effects of varying schemata on a reader's comprehension.

Miscues

Another relatively well developed arena of in-process indicators is miscue analysis. A miscue is an oral reading response that, in the judgment of an observer, deviates from the expected response (Goodman, 1965). We include the notion of the observer as a judge because miscue identification varies a bit depending upon the a priori schema, the dialect, the hearing, and other characteristics of the observer (Page, 1973). Miscue analysis represents a substantial change in the direction of oral reading analysis as it occurred for decades prior to Goodman's (1969) work. Weber (1968) describes prior work showing an emphasis on treating deviant oral reading responses as errors. Goodman (1969) demonstrates that miscues are errors only in relation to the task of producing a compliant phonological analogue of the author's printed language, and that if comprehension is considered to be the task, many miscues are indicators of the fact that the reader is understanding the author's message.

Miscue analysis procedure involves tape recording a youngster's oral reading and subsequent retelling of the stories. The procedure includes analyzing both the miscues and the retelling extensively, using a psycholinguistic taxonomy on the miscues and a technique of assessment of plot, theme, characters, and details to yield a comprehension rating of the retelling (Goodman, 1969). The creation of the miscue procedure has spawned a large number of studies during the past decade and a half, many of which are explained by Allen and Watson (1976).

A result of some of the miscue studies is the identification of a series of linkages between product and in-process indicators of reading comprehension, linkages that promise studies that can establish a stronger framework of corroborative indicators of reading comprehension. The creation of the comprehending score is an example. A comprehending score is the sum of percents of semantically acceptable miscues and successfully corrected semantically unacceptable miscues (Rousch, 1972, 1976). Goodman and Burke (1973) demonstrated that comprehending scores share significant variance with retelling ratings of passages that are read orally. Page (1977) exhibited a significant positive correlation between comprehending scores and post oral-reading cloze test scores. Carey (1978) found standardized reading comprehension test scores, post oral-reading cloze test scores, and comprehending scores positively and significantly correlated.

The linkages exhibited in studies with the comprehending score suggest new areas of investigation that are warranted in the efforts to form a stronger corroborative framework of comprehension indicators. The comprehending score consists of the sum of the percents of totally semantically acceptable miscues and totally semantically unacceptable miscues that are successfully corrected.

This finding suggests that variants of the comprehending score including combinations of percents of partially semantically acceptable miscues and partially corrected miscues may yield useful indexes of reading comprehension drawn from responses elicited during processing of printed language. One might speculate that several variants of the comprehending score are likely to correlate significantly with post oral-reading cloze scores and that some of these variants may distinguish reading ability groups. This approach is like a pincer movement, focusing on miscue types that correlate with comprehension as indicated by a product measure, and distinguish reading ability groups as indicated by teachers' observations of classroom performance of students in reading.

In another approach to creating combinatorial miscue type variables, studies can focus on miscue correction behavior and ratios of one type of miscue correction behavior to another. The percents of miscues that elicit successful, unsuccessful or no correction attempts may share significant amounts of variance with post reading comprehension measures. The percent of successfully corrected miscues divided by the percent of uncorrected unacceptable miscues may yield a ratio that correlates positively and significantly with post oral-reading cloze test scores. Similarly, the percent of successfully corrected miscues divided by the percent of all miscues that elicit no correction attempt may yeild a ratio that correlates significantly and positively with post oral-reading cloze test scores. Variables created out of combinations and ratios of observed correction behavior can be tested to determine if they distinguish ability groups. It is sufficient to note that a host of miscue combination variables have yet to be tested. Further studies probing relationships between these variables and comprehension, along with tests of how well these variables can distinguish reading ability indicated by classroom reading performance promise to provide information that can help to fill in major areas of the in-process portion of a web of corraborative indicators of comprehension.

The number of miscue types, the combinations, and the ratios together form a very complex set of possibilities. Page (1978a) describes the concept of an altercue continuum, a classification that focuses on the correlative relationship of each miscue type to other indicators of comprehension. A miscue is a deviation from the expected response if the expectation is a production of the sounds represented by the graphical display. If the expected response is formulated in terms of how meaningful the oral response is, then deviations from the phonological expectation may be thought of as alterations or altercues rather than miscues. Whether the deviations from the phonological response are called miscues or altercues, the physical data remain the same.

The concept of an altercue continuum is based on the fact that frequency counts of phonologically identifiable deviations from the expected phonological response can be cast in a bivariate relationship to some other indicator of comprehension, and that a correlation coefficient can be calculated in every instance. Because all miscue types, all combinations of miscue types, and all ratios of miscues that are actually produced will yield a correlation coefficient with a post reading comprehension indicator, a finite set of expectations exists. All Pearson product-moment correlation coefficients fall on a continuum from +1.00 to

−1.00; therefore, all miscues that are produced will exhibit a relationship to comprehension.

The concept of the altercue continuum provides a way of classifying all miscue types on the basis of a single unifying principle, the correlative relationship of the particular miscue type with comprehension. Three obvious mathematically determinable benchmarks reside on the altercue continuum. There are miscue types that exhibit high positive correlations, high negative correlations, and low correlations with comprehension. Page (1978b) labels miscues that exhibit high positive correlation coefficients with comprehension "supercues," miscues that exhibit high negative correlation coefficients with comprehension "pseudocues," and miscues that exhibit low or non-significant correlation coefficients with comprehension "entropicues."

Carey (1978) applies the concept of an altercue continuum to miscues classified according to correction behavior and semantic acceptability, categories drawn from the Goodman Taxonomy of Oral Reading Miscues (1969). Carey found five miscue types functioning as supercues in relation to standardized reading comprehension scores and cloze scores and three miscue types functioning as pseudocues. He suggests that miscues classified on the altercue continuum according to semantic acceptability, correction behavior, and how they relate to comprehension may provide a summary of how readers monitor the semantic aspects of their own processing.

Carey's concept of monitoring a reader's processing suggests that a model of monitoring during processing might be constructed to organize information drawn from the reader via miscues or other observable performance. Using miscues as an example, the order and sequence of miscue production can serve as one dimension of such a model. Each miscue occurs in a sequence that is directly related to the sequence in which the elements of the test are encountered. Some deviation occurs between the sequence of processing and the sequence of words as they appear printed in English text. Proficient silent readers may visually scan text in directions other than left to right. However, to a considerable degree, processing in oral reading occurs in the order that the text presents itself, perhaps because the oral reading task itself is usually conceived by readers as producing a spoken analogue of the printed text, word by word according to the sequence dictated by the text.

At each miscue, it is possible to produce a static representation of the characteristics of the reader's processing at that point in the sequence. In a model of monitoring, Page (1978c) calls the static representation a miscue frame, and compares it to a snapshot of a person running. A miscue frame can include a detailed description of the miscue. It might include descriptions of how the miscue relates to the reader's uses of semantic, syntactic, and grapho-phonemic cue systems as well as how readers use these same cue systems to overtly correct or not correct their miscues. The methodology of deriving descriptions of a miscue to fill out a miscue frame is well developed in the descriptive procedures provided in miscue research. Correction behavior is the key source of information about the reader's monitoring behavior of each miscue, and the miscue frame provides an organized bin for this information.

Much as moving pictures are created by producing a series of still pictures and showing them one after another in rapid sequence, miscue frames can be examined in the sequence that the reader produces the miscues. When miscue frames are viewed in sequence and in relationship to one another, it is possible to determine how readers monitor their own oral reading. The dynamic aspects of processing printed language can be portrayed as changes in the described characteristics of miscues shown in the miscue frames viewed through time. Shifts in the reader's mode of monitoring become evident, and can be shown as monitoring changes in relationship to the characteristics of text which trigger differing characteristics in miscue frames.

Both the cloze and miscue methodologies stand as reasonably well developed approaches to organizing overt indicators of comprehending drawn from the reader while the reader is processing printed language. We present ideas about cloze and miscue approaches here as examples, knowing full well that other avenues also warrant study. Both cloze and miscue techniques present difficulties because the processing of the reader under the prescribed conditions must differ to some degree from processing that takes place during ordinary silent reading. Both cloze and miscue techniques warrant further development. Models that can help organize the dynamic characteristics of processing are required if we are to more fully understand how reading works during processing.

Toward a Corroborative Framework

Both product and in-process indicators of reading comprehension present promising directions for further research. The concepts of product and in-process indicators suggest a dimension of a corroborative framework that can be developed by searching for and creating ways of eliciting and organizing information drawn from readers. The dimension itself suggests a unifying principle; the notion of when the overt responses are drawn from the reader in relationship to when the actual processing of the author's printed language occurs. Indicators of the reader's a priori schema may be elicited prior to encountering the text. In-process indicators may be elicited during the processing of the text. Post-process indicators may be elicited immediately following the reader's processing of the text. Delayed responses to the text may be elicited after varying amounts of time have passed after the processing of the text. Time in relation to the occurrence of the initial processing of the text represents a powerful conceptual dimension for organizing at least part of the web of corroborative indicators of reading comprehension.

Types of language units represent a second dimension applicable to constructing a corroborative web of indicators of reading comprehension. The dimension of types of language units yields to a single unifying principle that focuses on the size of the language unit. Benchmarks on the language unit size continuum might be characterized as small, medium, large, and very large, from microstructure to macrostructure. Small units of language involve letters, graphemes, phonemes, syllables, letter combinations, bound and free morphemes, and words. Medium sized units might include phrases, clauses, and sentences. Large size units might include paragraphs, passages, stories, and books; while very large units might be thought of as universes of discourse, whole disciplines, and the languages of various cultures or nations.

By casting the four benchmarks on the time continuum against the four benchmarks on the language unit continuum, a 4 x 4 structure with 16 cells comes into sharp focus. A priori schemata, indicators of processing during the initial reading, immediate post reading indicators, and delayed post reading indicators can be probed in terms of small, medium, large, and very large units of language. These two dimensions of a corroborative framework of indicators of reading comprehension are relatively easy to consider because each is intuitively clear. Other dimensions can be brought into the framework. For instance, the dynamic aspects of viewing reading during processing as it occurs through time can be superimposed on the in-process category of the framework.

There is little need here to expand on the possibilities of a corroborative framework of indicators of reading comprehension. We have provided a bold map of some of the more obvious possibilities, a map that can spawn further development. We have touched on the high points of the framework we conceive, and we have provided some examples from existing research of what might be included in such a framework.

The major development and use of a corroborative framework of indicators of reading comprehension is yet to come. In schools at present it is probably a rare occurrence when educators take into consideration the full range of possibilities of even the sparse framework we have described here. When we consider the implications of educators' decisions about how youngsters read, one can only marvel that so little information is used to control, and in fact form, the destinies of so many of our children. With mandated reading levels on the horizon for most states, one can make a reasonable case for the notion that researchers, test makers, and educators ought to be engaged in a fervent scramble to make sense of the less developed cells in a corroborative framework of indicators of reading comprehension. Perhaps the framework presented here can serve as a basic outline from which to work. Perhaps the notions we presented here can serve to trigger a better way of approaching the task.

References

Allen, P. D., & Watson, D. (Eds.). *Findings of research in miscue analysis: Classroom implications.* Urbana, Ill.: ERIC/CRCS, NCTE, 1976.

Anderson, R. C. How to construct achievement tests to assess comprehension. *Review of Educational Research,* 1972, *42*(2), 145-170.

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. H. Bower (Ed.), *The psychology of learning and motivation,* 9. New York: Academic Press, 1975.

Ausubel, D. P. *Educational psychology: A cognitive view.* New York: Holt, Rinehart, & Winston, 1968.

Bormuth, J. R. Factor validity of cloze tests as measures of reading comprehension. *Reading Research Quarterly*, 1969, *4*(3), 358-365.

Bormuth, J. *On the theory of achievement test items*. Chicago: University of Chicago Press, 1970.

Bormuth, J. Literacy in the classroom. In W. Page (Ed.), *Help for the reading teacher: New directions in research*. Urbana, Ill.: National Conference of Research on English, ERIC/CRCS, 1975.

Carey, R. F. *A psycholinguistic analysis of the effects of the semantic acceptability of oral reading miscues on reading comprehension*. Unpublished doctoral dissertation. University of Connecticut, 1978.

Carroll, J. B. Defining language comprehension: Some speculations. In R. Freedle & J. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston & Sons, 1972.

Farr, R. *Reading: What can be measured?* Newark, Del.: International Reading Association, 1969.

Frase, L. T. Prose processing. In G. H. Bower (Ed.), *The psychology of learning and motivation*, 9. New York: Academic Press, 1975.

Frederiksen, C. H. Acquisition of semantic information from discourse: Effects of repeated exposures. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 158-169.

Gagne, R. M., & White, R. T. Memory structures and learning outcomes. *Review of Educational Research*, 1978, *48*, 187-222.

Ganier, A. The post oral-reading cloze test: Does it really work? *Reading World*, 1976, *16*(1), 21-27.

Goodman, K. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, *42*, 639-643.

Goodman, K. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly*, 1969, *4*(1), 9-30.

Goodman, K., & Burke, C. *Theoretically based studies of patterns of miscues in oral reading performance*. U.S.O.E. Project No. OEG-0-9-320375-4269. Washington, D.C.: HEW, 1973.

Harper, R., & Kilarr, G. *Reading and the law*. Urbana, Ill.: International Reading Association and ERIC, 1978.

Horton, R. *The construct validity of cloze procedure: An exploratory factor. Analysis of cloze, paragraph reading and structure of intellect*. Unpublished doctoral dissertation. Hofstra University, New York, 1973.

Marshall, N., & Glock, M. D. Comprehension and connected discourse: A study into the relationship between the structure of text and information recalled. *Reading Research Quarterly*, 1978-79, *14*, 10-56.

Mayer, R. E. Information processing variables in learning to solve problems. *Review of Educational Research*, 1975, *45*, 525-541.

Page, W. D. Clinical uses of miscue research. In K. Goodman (Ed.), *Miscue analysis: Applications in reading instruction*. Urbana, Ill.: ERIC/CRCS, NCTE, 1973.

Page, W. D. The post oral reading cloze test: New link between oral reading and comprehension. *Journal of Reading Behavior*, Winter, 1975.

Page, W. D. Comprehending and cloze performance. *Reading World*, 1977, *17*(1), 17-21.

Page, W. D. The altercue continuum: Theoretical considerations. Abstract in *Resources in Education*. Reproduced in ERIC/CRCS. ED 155677, 1978. (a)

Page, W. D. Reading comprehension and oral reading. In R. T. Vacca & J. Meagher (Eds.), *Reading as a language experience*. Storrs, Conn.: Reading-Study Center, University of Connecticut, 1978. (b)

Page, W. D. *Models and reading*. Presentation at Indiana University, Bloomington, 1978. (c)

Page, W. D., & Vacca, R. T. *In-process and product indicators of reading comprehension: Toward a corroborative framework*. Paper presented at the National Reading Conference, St. Petersburg, 1978.

Pearson, P. D., & Johnson, D. *Teaching reading comprehension*. New York: Holt, Rinehart, & Winston, 1978.

Rickards, J. P. Interaction of position and conceptual level of adjunct retention of text. *Journal of Educational Psychology*, 1976, *68*(2), 210-217.

Rickards, J. P., & Hatcher, C. W. Interspersed meaningful learning questions as semantic cues for poor comprehenders. *Reading Research Quarterly*, 1977-78, *12*(4), 538-553.

Rothkopf, E. Z. Experiment on mathemagenic behavior and the technology of written instruction. In E. Z. Rothkopf & P. E. Johnson (Eds.), *Verbal learning research and the technology of written instruction*. New York: Columbia University Teachers College Press, 1971.

Rothkopf, E. Z. Structural features and control of processes in learning from written materials. In R. Freedle & J. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston & Sons, 1972.

Rousch, P. *A psycholinguistic investigation into oral and silent reading, prior conceptual knowledge and post reading performance: A study in miscue analysis*. Unpublished doctoral dissertation, Wayne State University, Detroit, Michigan, 1972.

Rousch, P. Testing. In P. D. Allen & D. Watson (Eds.), *Findings of research in miscue analysis: Classroom implications*. Urbana, Ill.: ERIC/CRCS, NIE, NCTE, 1976.

Smith, F. *Understanding reading*. New York: Holt, Rinehart, & Winston, 1971.

Taylor, W. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 1953, *30*, 415-433.

Tuinman, J. J. Determining the passage dependency of comprehension questions in five major tests. *Reading Research Quarterly*, 1973, *9*(3), 206-223.

Tyler, R. The use of tests in measuring the effectiveness of educational programs, methods, and instructional materials. In R. Tyler & R. Wolf (Eds.), *Crucial issues in testing*. Berkeley, Calif.: McCutchan, 1974.

Weber, R. The study of oral reading errors. *Reading Research Quarterly*, 1968, 4, 96-119.

William D. Page, Ed.D., is associate professor of education, University of Connecticut, Storrs. Richard T. Vacca, Ed.D., is associate professor of education, University of Connecticut, Storrs.

Cloze Restoration, Recall, and Discourse Processing in Middle Grade Readers

Sandra M. Gould

A major problem for researchers and educators interested in prose comprehension is how best to measure what is understood from a text. Assessment must be valid, reliable, and parsimonious; that is, the measure must be concomitantly accurate in what it purports to measure, predictive of subsequent performance, and inexpensive of time and other resources. The cloze procedure has been offered as an instrument that meets all three criteria. Cloze tests constructed and scored in the manner described by Bormuth (1975) claim to "measure a person's ability to perform the comprehension processes" by tapping these processes at two points: "testing how much knowledge was obtained from the text surrounding the blank and how well the information obtained from the text was employed to obtain additional information" (p. 75).

Earlier research conducted by Bormuth (1963, 1969) to study the relationship between cloze tests and traditional multiple choice comprehension tests is frequently cited, recently in Asher, Hymel, and Wigfield (1976), as evidence that cloze tests are valid measures of comprehension. Bormuth (1975) attests that the two kinds of tests are so highly correlated that "the processes that generate those scores are indistinguishable" (p. 67). His willingness to anchor cloze tests in this way is curious since Bormuth (1969) as well as others (Goodman, 1973) have elsewhere been critical of what is measured by standardized comprehension tests.

In contrast to the cloze procedure, which taps the contiguous production of text, recall assessments, taken subsequent to discourse processing, have also become accepted measures of comprehension. While researchers have adopted their own idiosyncratic variations, assessment generally employs one of two techniques: either free response such as paraphrasing or summarizing, or probe (cued) retrieval such as sentence recognition and questioning.

Free recall responses obviously create scoring problems butemerge as relatively uncontaminated estimates of the encoding process once objectivity is secured in protocol scoring. Free recall assessments have gained credibility in recent years with the development of story grammars that designate the specific unit of analysis (e.g., proposition, concept, or set relation) a priori and serve as a template for scoring protocols.

Two Views of Discourse Processing

Bormuth's (1975) description of what is involved in comprehending implies the *interpretist* model of discourse processing. Interpretists, who rely heavily on transformational and computational linguistics, assume that comprehension comes about through the automatic parsing of each sentence of input. This parsing, which computes the relationship between surface and deep structure, is presumed to correspond with the mental operations involved in understanding. That is, once a sentence is parsed, it is understood.

The problems associated with making any model fit the observations from discourse processing led to a growing disenchantment with the sufficiency of explanations which relied solely on semantic interpretations from deep structure to account for the full analysis of cognitive meaning. Increasingly, researchers began to examine a second model of discourse processing termed *constructivist.*

Broadly defined, the major difference between the approaches of interpretists and the constructivist is the disagreement on the locus of meaning. The interpretist's reliance on deep structure reflects the tacit assumption that sentences carry meaning. Constructivists argue that the individuals are the meaning carriers and linguistic inputs act only as cues by which they can recreate or modify their previous knowledge of the world (Bransford, Barclay, & Franks, 1972). The parsing process so critical to the interpretists' theory is another source of difference. Parsing suggests that sentences must be processed in their entirety. The constructivist's conception is that input is selectively processed, the text is sampled and combined with contextual information and stored knowledge of the world to produce an interpretation that fits the data. Comprehension from the constructivist's vantage involves using "prior knowledge, intentions, context and task demands in combination with input structure to control discourse processing" (Frederiksen, 1975, p. 140).

Rationale

Research carried out under the aegis of the constructivist model of discourse processing had broad implications for reading research, particularly in the measurement of comprehension. The task of adequately assessing comprehension continues to be a vexing problem. The parameters of concern may be defined this way: should comprehension measures focus on the more easily assessed *reproduction,* or iterative levels of processing; or should they center on the less easily appraised, but perhaps more accurate, *reconstructive* aspects. The latter position is supported in educational contexts where teachers are interested in students' abilities to apply understandings from reading to new situations.

The theoretical framework for this research was drawn from the constructivist position that more is involved in the comprehension of discourse than reduction of semantic and syntactic sentence structure to their deep level constituents. This interpretist model of information processing, which has acted as the justification of the cloze technique as a comprehension measure, has been viewed as less than adequate in the face of increased research in processing beyond the level of the sentence.

Many theorists have now come to believe that subjects encode sentences neither in their entirety nor as kernels. Instead they are viewed as engaging in an integrative process in which the ideas of the message are incorporated into the cognitive structure in such a way that the end result is a whole no longer separable into summative parts. Jenkins' (1974) discussion of the quality of events exemplifies this synergistic view of the constructivist process.

The position that this constructive process is influenced not only by prior knowledge and input content and structure but also by context and task demands gave rise to a sub-paradigm within the constructivist camp labeled "levels of processing." Craik and Lockhart (1972) initiated investigation into this alternate framework for memory research which has developed into a compelling exemplar for studies of comprehension and memory for printed stimuli. Essentially, the depths of processing model argues that memory traces are automatic by-products of operations carried out by the cognitive system and the trace's strength is a "positive function of 'depth' of processing where depth refers to greater degree of semantic involvement" (Craik & Tulving, 1975, p. 268). Studies carried out under incidental as well as intentional learning conditions (Craik & Tulving, 1975; Frederiksen, 1975; Jenkins, 1974) make a powerful case for the view that it is the operation carried out at the time of encoding and *not* the intention to learn that is the prime determiner of retention.

Recently new evidence has emerged to suggest that the original depth of processing stance needs refinement. Nelson (1977) argues that the position lacks credibility in the absence of an algorithm for ordering different kinds of processing in terms of depth. At present ordering is circular since the "depth" of various processing tasks has been determined after its effects on memory were known.

Morris, Bransford, and Franks (1977) have suggested that the concept of "levels" should be replaced by "transfers appropriate processing." They propose that tasks are neither inherently superficial nor meaningful; rather, their value must be defined relative to the particular goals and purposes for which they were intended. Further, the original assumption that depth was intrinsic in the task ignores the readiness state of the processor. In order to remember well, subjects must be able to use their knowledge and skill to process input.

These arguments underline the importance of carefully defining those factors that can operate to prevent processing from reaching deep elaborative levels. Previously, Craik and Lockhart (1972) had outlined three sources of failure: the nature of the material, limited available processing capacity, and task demands. The "transfer appropriate processing" position suggests another barrier might be a lack of congruence with intended learning goals. That is, the task set at the time of processing may be irrelevant, or counter-productive, to the desired learning outcome. A fifth obstacle, according to Spiro (1977), may be the context of the experimental setting which increases the likelihood that the subjects will not integrate the stimuli with their pre-existing cognitive structure; that is, experimental subjects tend to believe that the examiner wants "pure" effects unconfounded by idiosyncratic knowledge.

Considerable research has accumulated to predict the comparative effects of orienting tasks requiring attention to either phonemic, graphemic, structural or semantic features. Little has been done to specify any differential outcomes within the same domain. A question which might be asked is, are there predictable differences in memory for tasks that presumably require some semantic processing (e.g., cloze restoration or instructions to perform logical ordering)?

Craik and Lockhart (1972) implied that depth might be ordered by degree of semantic analysis, but they offered no specifications for determination. Craik and Tulving (1975) suggested that "spread" was a better metaphor but still did not suggest any method of ordering. Work completed recently by Postman and Kruesi (1977) may offer a means for resolving this problem. They argue that the systematic differences in recall within a domain may be the function of what they call "displaced rehearsal." Orienting tasks that require the subject to establish inter-item relationships will encourage rehearsal of the items and will produce greater recall. This is predicted whether the domain is semantic or phonemic.

Clearly, it is the semantic domain that is of primary concern to the goal of discourse understanding. If Postman and Kruesi (1977) are correct, then any task known to require inter-item integration should produce superior recall performance over a task that does not. Likewise any task known to require integration of information across sentence boundaries should produce greater recall than a task whose requirements are limited largely to within-sentence integration.

This research was intended to provide empirical demonstration of this prediction. Recent research on levels of processing suggests that within the semantic domain, orienting tasks can produce systematic differences in recall depending on the degree of elaborative or integrative processing that is required. It seems clear that the depth at which processing operates is heavily dependent on the usefulness to the subject relative to the task required of him. It follows that if an orienting task can be accomplished without integrating information across sentence boundaries, the subject will not be inclined to do so. This reluctance is particularly more likely in experimental settings, according to Spiro (1977). While successful completion of cloze units does require attention to within-sentence semantic and syntactic constraints, the demonstrated relationship to redundancy utilization (Carroll, 1972; Horton, 1975; Tuinman, Blanton, & Gray, 1975) is largely irrelevant to acquisition of major propositions that require integration across sentences.

While there is sufficient literature to make firm predictions about the effects of a task on recall of explicit propositions, a task's effect on inference production is not so readily anticipated. The appearance of inferences in recall protocols has not been examined sufficiently to warrant strong predictions on what generates or prevents them. Spiro (1977) has suggested that inferences, whether accurate or inaccurate, may be the result of either active elaborative processing at encoding or merely bad guessing to fill the gaps in memory. It follows then that given this experiment's instructions to recall everything as close to verbatim as possible, the subject would be discouraged from including inferences in free recall even though they were available. A demonstrated negative relationship between explicit and implied propositions would suggest that the incidence of inferences, under these conditions, corresponds to the need to fill gaps in memory. When instructions demand recall as accurately as possible, the absence of inferences in free recall could not be interpreted to mean that subjects did not process any inferences; they simply chose not to report any. The cued recognition test of true and false propositions was intended to assess the true availability of inferred propositions as a consequence of task.

Method

Rather than accept criterion cloze restoration as cannonical proof of comprehension, this research employed the cloze technique as one orienting task whose effects on recall were studied relative to another production task under both incidental and intentional conditions. If it is true that cloze tasks measure a sensitivity to local redundancy and not the comprehension of main ideas, then cloze restoration should not be facilitative of incremental gains in proposition recall when compared with a production task which requires performance of logical ordering on inter-sentential semantic relations. This prediction is reinforced by the levels of processing position; what is recalled is a positive function of operations carried out at the time of encoding and greater semantic recall arises from greater degree of elaborative or integrative coding of the text at the time of input.

If these predictions are correct, this would suggest that the mental operations being tapped by the cloze technique are not sufficiently inclusive to justify the cloze procedure as a measure of the comprehension processes.

Hypotheses

Three major hypotheses were of interest in this experiment.

Hypothesis 1. If the constructivist position is correct, that is, if parsing of sentences is not sufficient to account for comprehension and recall, then instructions to perform logical ordering across sentences (list), should result in significantly greater free recall of explicit propositions than cloze tasks performed under either incidental or intentional conditions ($\text{Group Three}_e > \text{Group Two}_e \geqslant \text{Group One}_e$).

Hypothesis 2. If it is true that cued recognition is more facilitative of detecting inferences that are available, then the elaborative coding predicted by the logical ordering task (list) should result in significantly greater correct recognition of true and false inferences than the automatic parsing carried out by both cloze conditions ($\text{Group Three}_{cri} > \text{Group Two}_{cri} \geqslant \text{Group One}_{cri}$).

Hypothesis 3. If it is true that instructions to report as close to verbatim as possible discourage the appearance of inferences except to fill memory gaps during free recall, then a negative relationship between explicit and inferred recall scores should be observed.

Population, Sample, and Sample Selection

The population for this study was composed of male and female children between the ages of 10 and 12 who were in grade 6 in a middle-class suburban school district west of Albany, N.Y. This population was restricted to those students whose parents had given permission for participation in the study. Of the 120 parents contacted, 76 parents granted permission. A total sample of 69 sixth-grade students was analyzed. Subjects were randomly selected and assigned to condition with an attempt to balance each group by sex.

Materials

All subjects were exposed to the same story passage of approximately 280 words which is a slightly altered version of a story titled *The Clever Turtle,* an African folktake adapted by A. K. Roche for Scott, Foresman Reading Systems, Level 6, Book B. Level 6 materials of this series are appropriate for use with average second grade youngsters. A natural story was chosen to guard against possible contamination from unique processing strategies invoked by artificially constructed discourse (Bower, 1976; Mandler & Johnson, 1977; Thorndyke, 1975).

Procedure

Students were tested outside their classrooms in one of three rooms depending on experimental condition. Three assistants, not necessarily naive to the study's purposes, read the written instructions aloud and monitored the written procedures. Students were told to write their names on all test materials. Recall and recognition protocols were later scrambled by the assistants and scored blind by two judges. Students were asked to read the passage under one of three experimental conditions: (a) instructions to restore cloze units with instructions to recall (Group One); (b) instructions to restore cloze units without instructions to recall (Group Two); or (c) instructions to read and list in sequential order the four things that the villagers wanted to do with the turtle with instructions to recall (Group Three). Processing time was controlled across conditions and was arrived at by taking the mean time required to complete the cloze task by a separate group of sixth-grade students. When that time had elapsed, all materials were collected, blank paper was distributed and students were asked to write all that they could remember of the story using exact words and in the exact order whenever possible. Students were told not to leave anything out just because they did not recall exact wording or order. The time for recall was unlimited but all were finished in less than 15 minutes.

When all students had finished, they were again reminded to check to see that their names were on the papers. The papers were collected and the cued recognition test was distributed. After recording their names, students were asked to respond to 12 propositions that might be inferred from the story by circling either True or False. Seven of these statements were true inferences while the remaining five had one or more words that made them false inferences. For example, the statement that "The turtle was not afraid of holes or rope" was a false inference although the villagers thought it to be true. Likewise, the statement that "The turtle wanted to be thrown into the river" would be a true inference although the turtle had stated just the opposite. All students were finished with the cued recognition test in less than five mintues.

Design

The researchers examined the relative effects of task variables on sixth-grade students recall and recognition of explicit and implied propositions and also studied the effects of cloze tasks under incidental and intentional conditions. A 3x3 design was employed with repeated measures over all conditions having one within-subjects variable (three production tasks) and three between subjects variables (explicit recall, inferred recall, and cued recognition). Analysis of variance and subsequent planned comparisons were employed to evaluate the significance of differences in recall and recognition trends for task conditions.

To facilitate scoring, the "Turtle" passage used in this experiment was segmented into 51 explicit propositions (a clause or sentence containing an action or stative verb). These propositions were then arranged in a tree structure according to Mandler and Johnson's (1977) rewrite rules. Scoring was by gist recall of the propositions based on a proper reconstruction of the action or relationship given. Synonymous paraphrase was permitted as was deletion of adjectives and adverb modifiers.

In addition to this story grammar which accounts for the story's explicit propositions, a list of 20 statements that might be legimately inferred from the passage was drawn up by two judges. This list was used to score protocols for incidence of inferred information. This list was also used to construct the cued recognition test of seven true and five false inferences. Scoring of the cued recognition test was according to the total number of correctly recognized true and false inferences. Scoring was carried out blind by two judges experienced in scoring proposition protocols. Differences were rare and were most frequently caused by failure to check off all propositions. Ease of scoring was facilitated by a general inclination to word-for-word reporting.

Results

An overall picture of the data is provided in Table 1 which contains a summary of the mean scores and standard deviations for all three dependent variables. One-way analysis of variance and planned comparison data indicated that the effect of task at this sixth-grade level did elicit significantly different outcomes for both explicit and inferred recall but made no observable difference in the case of cued recognition of inferences.

Hypothesis 1 predicted that instructions to perform logical ordering across sentences (list) would result in significantly greater free recall of explicit propositions than cloze tasks performed under either incidental or intentional conditions. This first hypothesis suggests that more is involved in the comprehension of discourse than the automatic parsing of sentences required by a cloze restoration task. This prediction was strongly supported. Analysis of variance (see Table 2) comparing the effect of task on recall of explicit propositions found a significant difference between the three task conditions, $F(2,62)=7.94$, $p<.001$. A subsequent orthogonal contrast comparing the list task with the two cloze tasks was also significant $t(62)=-3.78$, $p<.001$, indicating that instructions to perform logical ordering were indeed more facilitative of explicit proposition recall than either cloze task.

Table 1
Mean Scores and Standard Deviations for Explicit Recall, Inferred Recall, and Cued Recognition as a Function of Task

	Explicit Recall	Inferred Recall	Cued Recognition
Cloze (intentional)			
(N=20)	17.75	1.25	8.9
	9.16 sd	1.48 sd	2.13 sd
Cloze (incidental)			
(N=25)	24.72	.44	8.5
	11.02 sd	.77 sd	2.3 sd
Logical Ordering (list)			
(N=20)	29.50	.70	8.7
	7.1 sd	.66 sd	2.8 sd

Table 2
Analysis of Variance—Task x Explicit Recall

Source	df	SS	MS	F
Groups 1-3	2	1399.08	699.54	7.94[a]
Error	62	5465.79	88.16	

[a] $p < .001$

Hypothesis 2 stated that if cued recognition was more facilitative of detecting inferences that were available, then the elaborative coding predicted for the logical ordering task (list) should result in significantly greater correct recognition of true and false inferences than either cloze condition. Analysis of variance (see Table 3) comparing the effect of task on cued recognition failed to support this prediction $F\ (2,62) = .138$, $p > .05$. An examination of the group mean scores and standard deviations (shown in Table 1) suggests that the testing instrument lacked discriminatory power and was possibly not capable of detecting differences if they were there.

Hypothesis 3 examined the supposition that examiner instruction to recall as accurately as possible would discourage the appearance of inferences in recall protocols except in cases where gaps in memory required it. If this were true, a negative relationship between explicit and inferred propositions would be observed. This supposition received partial support from two sources. Analysis of variance (see Table 4) studying the effects of task on inferred recall $F\ (2,62) = 3.59$, $p < .05$, revealed that the groups were different. Examination of the mean scores shown in Table 1 leads one to conclude that it was the group scoring lowest in explicit recall, intentional cloze, that was responsible for the significance of the difference found in inferred recall. Additionally, Pearson correlational coefficients calculated for explicit and inferred recall demonstrated a negative trend in all three tasks (−.13, −.03, −.31) although none was significant (Table 5).

Table 3
Analysis of Variance—Task x Cued Recognition

Source	df	SS	MS	F
Groups 1-3	2	1.63	.81	.138[a]
Error	62	364.59	5.88	

[a] $p > .05$

Table 4
Analysis of Variance—Task x Inferred Recall

Source	df	SS	MS	F
Groups 1-3	2	7.43	3.71	3.59[a]
Error	62	64.11	1.03	

[a] $p < .05$

Table 5
Summary of Pearson Correlation Coefficients for Explicit and Inferred Recall, and Inferred and Emphatic Ending

	Cloze Intentional	Cloze Incidental	List
Explicit/Inferred	-.13	-.03	-.31
Inferred/Ending	.45[a]	.82[b]	.52[a]

[a] $p < .05$
[b] $p < .001$

Discussion

Of the three hypotheses associated with this experiment, only the first was strongly supported statistically. Instructions to perform logical ordering across sentences was demonstrated to be superior to either cloze task on the explicit recall measure. This finding would suggest that the constructivist position regarding automatic parsing of sentences is correct. Successful completion of a parsing task, i.e., cloze restoration, was not sufficient to explain the full range of comprehension processes. Students required to list in sequence the four major events of the story were better able to recall the story's explicit content. This was true despite the fact that while cloze restoration required an average of 15 minutes for completion, most students completing the list task were finished in less than 5 minutes. This finding reinforces Craik and Tulving's (1975) earlier conclusion that processing time cannot be reliably used as an indicator of depth. However, it does provide some additional support for the hypothesis that tasks within the semantic domain might be ordered by degree of inter-item integration required (Postman & Kruesi, 1977).

The failure to obtain support for the prediction of Hypothesis 2 (that cued recognition would detect an incremental gain in inferences processed by subjects engaged in the more elaborative coding required by the list task) was likely a result of the testing instrument. The use of a true/false paradigm using inferences that might legitimately be drawn from the story had been justified by its use in other studies of inferences by children (Paris, 1975). However, it seems clear that this particular instrument lacked discriminatory power. In retrospect, Mosher's (personal communication, February 1978) recommendation of the use of a forced choice recognition paradigm would seem more prudent.

It is duly noted that subjects in all three task conditions may have processed inferences similarly. Subjects in all three conditions did, after all, score above chance levels on the cued recognition task. In this case, it is well to remember Postman and Kruesi's (1977) admonition that the imposition of an orienting task does not give the experimenter full control over the subject's encoding activities. In fact, the familiar theme that underlies the Clever Turtle story may well have made it easy to recognize inferences. No apology need be made for recognizing that the orienting task does not totally control how the text is encoded; it is just a reaffirmation of the constructivist position that a number of variables interact to elicit textual understanding.

The experiment's prediction in Hypothesis 3 (that examiner instructions to recall accurately would discourage the appearance of inferences in recall protocols which, in turn, would encourage their appearance only as a filler to memory gaps) received some support from two sources. First, it was the group that scored lowest in explicit recall that was responsible for the significant main effect of difference in inferred recall. Second, the correlations between explicit and inferred recall were negative for all three groups.

Examiner observations as well as the recall protocols gave credence to Spiro's (1977) suggestion that experimental subjects tend to give back pure effects uncontaminated by idiosyncratic knowledge. Both judges noted that scoring of

78

protocols was made easy by the pronounced tendency of subjects to report word for word when they could do so. Two examiners noted that students had requested a restatement of directions to be sure the examiner did not want a summary.

Inspection of the means and standard deviations for inferred recall (Table 1) clearly reveals that inferences appeared rarely in the student recalls. Where they did appear, they were significantly correlated with the addition of an emphatic ending (Table 5). In other words, subjects in all conditions showed a tendency to add a more conclusive ending to the story such as "and the turtle got away" or "the turtle lived happily ever after." This tendency, which also has been noted elsewhere (Mandler & Johnson, 1977; Spiro, 1977) suggests that the appearance of inferences in subjects' recall cannot automatically be assumed to reflect deeper processing. Inferences may simply indicate items from script or knowledge of the world filled in by default because the subject either did not encode the full message or could not remember it (Kuipers, 1975; Pearson & Nicholson, 1976).

Summary and Pedagogical Implications

Arguing from a constructivist position, this research postulated that more is involved in the comprehension of discourse than the automatic parsing of sentences required by cloze tasks, and that cloze tasks would be demonstrated to be relatively shallow processing tasks when compared with tasks requiring intersentence integration. This position was supported by the results on recall of explicit propositions but no differences were demonstrated for cued recognition. The results reported for inferred recall tend to support, although tentatively, the predictions that (1) experimental context can adversely affect production and reporting of inferences in recall; and (2) inferences, in such cases, may reflect memory gaps rather than elaborative processing.

The outcomes of this research clearly indicate that cloze restoration to criterion is not equivalent to recall of textual propositions although both share the label "comprehension." This finding is important since it underlines the danger of generalizing across operational definitions of the construct called comprehension; and it suggests that care must be taken in making choices among the comprehension measures that are available.

Selection of assessment measures for textual understanding continues to be a major problem in education. In the past, the search for testing instruments emphasized three qualities: validity, reliability, and parsimony. That cloze procedures meet all three criteria is not challenged by this research. Cloze tests do appear to measure what they claim to measure, scores are predictive of subsequent cloze testing, and test construction is easily and inexpensively done. What is at issue is the sufficiency of these qualities for all educational applications.

This research suggests that a fourth demand be made of measuring devices—*suitability*. That is, before selecting a comprehension test, it is important to examine the linkage between what the instrument purports to measure, i.e., how comprehension has been operationalized, and what learning outcomes are desired. Cloze tests claim to tap the comprehension process by "testing how much knowledge was obtained from the text surrounding the blank and how well the information obtained from the text was employed to obtain additional information" (Bormuth, 1975, p. 67). By definition, cloze tests emphasize the reproductive aspects of understanding; students are required to restore cloze units verbatim. Test instructions openly discourage reconstruction and, consequently, encourage the belief, noted by Spiro (1977), that the examiner wants "pure" effects unconfounded by interaction with cognitive structure.

It is true that recall measures can also emphasize reproduction without cognitive integration. However, unlike cloze tests, it seems possible that recall measures can be designed to enhance reconstruction. Given story grammars that include appropriate instructions and an appropriate range of implicit as well as explicit propositions, students can be encouraged to decompartmentalize text information.

If one educational priority is to maximize the integration of previous knowledge with information gained from a text, then it would seem essential that assessment devices reflect this goal. Clearly, comprehension assessments must be valid and reliable. But they also must be appropriate to learning goals and this suitability must not be superseded by economy.

References

Asher, S., Hymel, S., & Wigfield, A. *Children's comprehension of high and low interest material and a comparison of two cloze scoring methods* (Tech. Rep. No. 17). Urbana-Champaign, Ill.: University of Illinois, Center for the Study of Reading, 1976.

Bormuth, J. R. Cloze as a measure of readability. In J. A. Figurel (Ed.), *Reading as an intellectual activity.* Newark, Del.: International Reading Association, 1963.

Bormuth, J. R. Factor validity of cloze tests as measures of reading comprehension ability. *Reading Research Quarterly*, 1969, *4*(3), 358-365.

Bormuth, J. R. The cloze procedure: Literacy in the classroom. In W. Page (Ed.), *Help for the reading teacher.* Urbana, Ill.: National Conference on Research in English, 1975.

Bower, G. Experiments on story understanding and recall. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 511-534.

Bransford, J., Barclay, J. R., & Franks, J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 1972, *3*, 193-209.

Carroll, J. B. Defining language comprehension. In J. Carroll & R. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* New York: Wiley & Sons, 1972.

80

Craik, F., & Lockhart, R. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671-684.

Craik, F., & Tulving, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, *104*(3), 268-294.

Frederiksen, C. Effects of context induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, *7*, 139-166.

Goodman, K. Testing in reading: A General critique. In R. Ruddell (Ed.), *Accountability and reading instruction*. Urbana, Ill.: National Conference of Teachers of English, 1973.

Horton, R. J. Construct validity of the cloze procedure. *Reading Research Quarterly*, 1975, *10*(2), 248-251.

Jenkins, J. Remember that old theory of memory? Well forget it! *American Psychologist*, November 1974, 785-795.

Kuipers, B. A frame for frames: Representing knowledge for recognition. In D. Bobrow & A. Collins (Eds.), *Representation and understanding*. New York: Academic Press, 1975.

Mandler, J., & Johnson, N. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

Morris, C. D., Bransford, J. D., & Franks, J. Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 151-171.

Paris, S. Integration and inference in children's comprehension and memory. In F. Restle (Ed.), *Cognitive theory* (Vol. I). Hillsdale, N.J.: Erlbaum, 1975.

Pearson, P. D., & Nicholson, T. *Scripts, texts and questions*. Paper presented at the National Reading Conference, Atlanta, Ga., 1976.

Postman, L., & Kruesi, E. The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 353-369.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Thorndyke, P. *Cognitive structures in human story comprehension and memory*. Santa Monica, Calif.: Rand, 1975.

Tuinman, J., Blanton, W., & Gray, G. A note on cloze as a measure of comprehension. *The Journal of Psychology*, 1975, *90*, 159-162.

Sandra M. Gould, Ed.D., is at Guilderland Central Schools, Guilderland, New York.

Measures of Comprehension: Independence and Implications

Pat Rigg and Elizabeth Taylor

Definitions and measures of comprehension vary widely. Despite these differences, theorists and practitioners generally agree on two points: (1) the goal of reading is comprehension; (2) one indicator, however imperfect, of comprehension is the ability to relate to someone else what one has read. This relating—or retelling, as K. Goodman and Burke (1973) call it—may also be a means of increasing comprehension, since the act of retelling requires organizing the information from the reading into a coherent form for presenting to another. Smith, Goodman, and Meredith (1976) suggest that this presenting phase is not just helpful, but is a necessary step in the long term integration of new information with old. Telling someone about what one has read requires much more than recollection; it requires reflection and it requires an organization of what is recollected into a sequence that forms an organized whole. Furthermore, if the speaker is at all sensitive to the audience it requires adjustment to that audience.

Retelling has been used by K. Goodman and his colleagues as a method of measuring comprehension for several years. Retelling is used for both research and classroom instruction. As a research tool, it is invaluable in eliciting whole, natural language focused on what the subject deems important. As an instructional technique, it focuses the student's attention on getting meaning from text at the same time that it allows the opportunity to process that meaning through recollection and reflection.

Retelling has two parts. First, the reader is asked to tell everything he or she can remember about the story. During this first, unaided, part of the retelling, the researcher acts as an interested listener, perhaps nodding occasionally, but never interrupting. When the reader makes it clear that he or she has nothing further to add, the researcher begins the second part of the retelling, during which the researcher asks questions based on what the reader has already mentioned. These questions ask for clarification and expansion of information already volunteered. The researcher cannot say, for example, "You forgot to mention so-and-so. Who was he?" The researcher can only ask, "Can you tell me anything else about the people in the story?"

Instructions for eliciting a retelling (K. Goodman & Burke, 1973) indicate that the retelling follows an oral, tape-recorded reading of a complete text. That is, the reader retells to the same person who has just heard him or her orally read the text. What effect this pattern has on the retelling is unknown. Perhaps it has little; perhaps schoolchildren are so accustomed to non-communicative uses of language that telling someone a story they have just read to him is as reasonable and acceptable as, say, reciting the spelling of a list of words. It is probably more reasonable since relating a story clearly is a great deal more meaningful than reciting either letters or unrelated words. To increase the meaningfulness of the retelling task in this study, the students read silently rather than orally. Their retellings, therefore, were not a repetition of the same information in another form, but real communication. The scoring system for retelling allows for comparing individuals and it suggests, in an admittedly incomplete way, how much of the reading material was comprehended.

Cloze is another indicator of comprehension. It is much more commonly used than retelling for two reasons: cloze is faster and easier to administer and to score than retelling; cloze correlates well with standardized measures (Bormuth, 1975). The rationale for using cloze as a measure of comprehension is that readers cannot satisfactorily fill in the blanks unless they understand the material. How much of a subject's correct score is due to the ability to manipulate the surface features of the language, and how much to real understanding of the text, is a debatable question.

A third measure of comprehension is exemplified by the California Achievement Test, which asks multiple-choice questions about each short reading passage. The CAT, like many standardized tests, is used by teachers and school administrators to group students, and is often used by researchers for the same purpose. In this study, CAT scores were used to assign subjects to a group, then one measure of comprehension (retelling) was manipulated, and the results were compared to cloze scores.

Research Questions

There were several questions the study was designed to answer. The first general question was, does comprehension as measured by retelling relate to comprehension as measured by cloze? More specifically, does the number of retellings—none, one, or two—affect the cloze score? The second general question was, does comprehension as measured by retelling change as retelling is manipulated? This second question was rephrased as two more specific questions: (1) does delaying the retelling one week significantly affect the retelling score; and (2) does retelling immediately after reading affect the score of retellings produced one week later?

Procedure

In order to answer those questions, 48 fifth-grade students from a rural-suburban school were ranked according to their CAT scores, then randomly assigned to four groups. That is, of the four highest-ranking students, one was assigned to Group NR (No Retelling), one to Group DR (Delayed Retelling), one to Group IR (Immediate Retelling), and one to Group 2R (Two Retellings). The next rank of four was assigned one to each group, and so on. Although the original list of subjects totaled 48, absences on one of the four data-collection days and mechanical problems with two tape-recorders reduced the number of subjects finally used in the study to 40.

All of the subjects read "My Brother is a Genius" (Betts & Welch, 1963), a selection of approximately 2,300 words. Groups IR and NR read on a Monday, Groups DR and 2R on a Friday. All groups read from 10 to 10:30 a.m. All subjects were asked to read silently without assistance, and they were told that they would be asked to write something about the story in a week. The 30 minutes allowed for reading the story was ample time; students who finished earlier (and all did) were given graph paper and a pencil and asked to design their name in the squares. Groups IR and 2R were told in advance of their reading that they would be asked to tell everything they could remember about the story after they finished reading. Groups DR and 2R were told in advance that they would be asked in a week to tell everything they could remember. The retellings were guided by persons trained in the administration of the Reading Miscue Inventory (Y. Goodman & Burke, 1972) and were tape-recorded for later analysis. One week later, Groups DR and 2R retold the story, and all groups received the same cloze text. The cloze passage was taken from the last page of the story. Every fifth word was deleted for a total of 50 blanks. Treatment by group is shown in Table 1.

Table 1
Order of Reader's Task by Group

NR	IR	DR	2R	Time
No Retell	Immediate Retell	Delayed Retell	Two Retells	
Read	Read	Read	Read	30 min.
	Retell		Retell	15 min.
		Retell	Retell	
Cloze	Cloze	Cloze	Cloze	15 min. One week
				15 min. later

Analysis of Data

The retellings were transcribed, and the transcriptions checked by an independent listener for accuracy. The transcribed retellings were then scored according to directions in the Reading Miscue Inventory (RMI). The RMI assigns points for different categories: recall of characters, recall of plot events, statement of theme, and statement of plot. The scoring of the retellings was done independently by two researchers trained in administering and scoring the RMI; their scores were compared and differences were resolved through discussion and, sometimes, consultation with a third researcher.

The cloze passages were scored twice. First they were scored for exact replacement of the original word; second, they were scored by two researchers independently for semantic acceptability of response. Semantic acceptability was defined according to RMI guidelines, so that any word or phrase or blank left blank which resulted in a sentence that made sense with the rest of the story, was acceptable.

One of the research questions was, does the number of retellings affect the cloze scores? In order to answer this, a one-way analysis of variance was run on the cloze tests, first using the exact replacement scores, and then using the semantically acceptable scores. In both cases, the four groups were not significantly different from each other. For the exact replacement scores, F=.366, (df=3,36), p<.778; for the semantically acceptable scores, F=.668, (df=3,36), p<.577. The means of the four groups on each cloze scoring are indicated in Figure 1. A slight linear trend is apparent, but it is not great enough to draw any conclusions

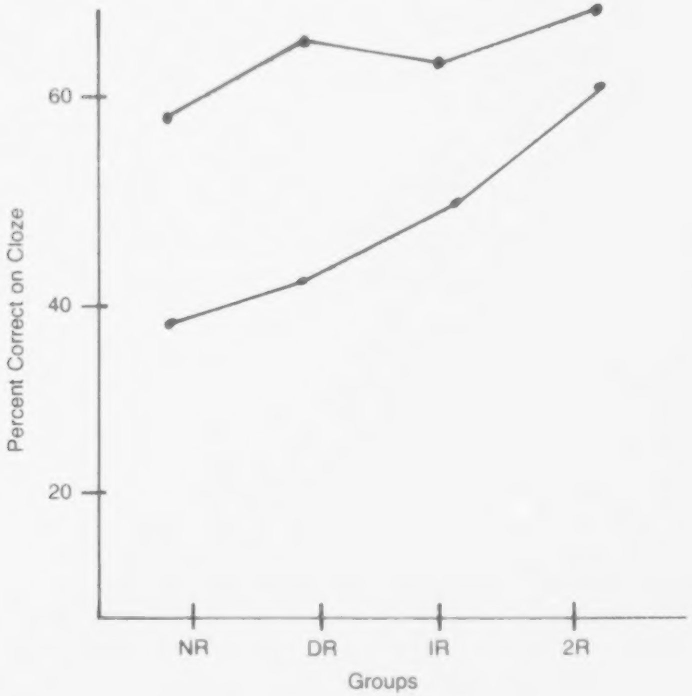


Figure 1. Cloze performance by group.

about the performance of the groups. Essentially, the groups did not differ on the cloze measure.

A second question was, does it make any difference in the cloze score if the retelling is immediate or delayed a week? As one may see in Figure 1, it makes none. The group that retold immediately (IR) had a mean of 44% exact replacement; the group that retold one week later (DR) had a mean of 41%, a nonsignificant difference.

All the results of the cloze tests in this study, both exact replacement scores and semantically acceptable alternative scores have little or no relationship to the number and timing of the retellings. The implication for the classroom is clear: whatever retelling may do for the students' comprehension, it cannot be expected to raise their scores on cloze tests.

For most of the groups, the retelling scores were significantly correlated with the cloze scores indicating that if a subject scored highly on the retelling task, he or she was also likely to score highly on the cloze task. However, this relationship did not hold for the second retelling scores of the group that retold the story twice.

A third question asked in this study was, does retelling immediately after reading affect the scores of retelling one week later? To measure the effect of retelling twice on the score of the second retelling, Group 2R had its first retellings compared to those of IR group, and its second retellings compared to those of the DR group. A one-way analysis of variance was used to compare these retelling scores. When the two groups that retold the second week were compared to each other, their scores differed significantly. $F=13.867$, $(df=1,18)$, $p>.01$. The mean scores of the retellings are shown by group in Figure 2.

We cannot know how much of this result is attributable to retelling immediately after reading, and how much is attributable to knowing that the immediate retelling will be followed by another retelling a week later. Clearly, just informing the students that they will be asked to retell next week does not elicit the same quality of retellings as informing them of this at the same time one asks for a retelling. One cannot assume, though, that the process of retelling immediately is solely responsible for the dramatic results.

The results of this study have several implications for the classroom. The first is that teachers who want their students to recall what they have red s right after reading it, and then again several days later. This is not to say that the teacher should administer an RMI-sort of retelling to each student after each bit of reading; such a practice is, of course, impossible. But students should be allowed, even urged, to tell someone, possibly a classmate, about what they have just read, and they should be given the time and encouragement to repeat this process in a week or so. It is not difficult to train students to be the sort of listener an RMI-administrator is —one who listens quietly and supportively, and who asks questions of clarification only. The teacher can model the sort of listening and questioning he wants, and then the students can pair off, reciprocally retelling and listening.

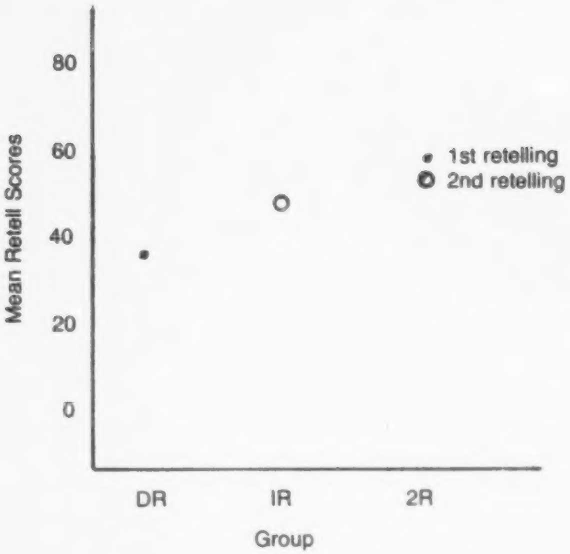


Figure 2. Mean scores of the retellings.

The second implication of this study for the classroom is that standard measures of comprehension should be used very warily. Obviously, the CAT and tests like the CAT which ask specific questions about short passages are measuring some sort of comprehension. The cloze test, too, is measuring some sort of comprehension. They may be measuring the same thing, but they are not measuring those aspects of comprehension that are elicited only by retelling.

A basic difference between the CAT, cloze, and retelling is that both the CAT and cloze ask specific questions about the reading material; in retelling the opening question is holistic: "Tell me everything you can remember." The CAT and a cloze test say to the reader, "This is important. What's the bit of information that fits this question?" Retelling says, "What do you think is important? How can you communicate that to someone?" The CAT and cloze measure what readers can produce that satisfies the requirements of the questioner; retelling measures what the readers can produce that satisfies themselves. This is a difference that cannot be emphasized too strongly: unlike the other tests, retelling asks the reader to establish the criteria of importance and of completeness. And, because retelling requires the reader to recall, reflect, and reorder the information from the reading material, it may very likely require a "deeper" cognitive processing than that necessary to answer CAT-like questions or to fill in cloze blanks.

References

Betts, E. A., & Welch, C. M. *Adventures now and then.* New York: American Book, 1963.

Bormuth, J. R. Literacy in the classroom. In W. D. Page (Ed.), *Help for the reading teacher.* Urbana, Ill.: NCTE/ERIC, 1975.

Goodman, K. S., & Burke, C. *Final report: Theoretically based studies of patterns of miscues in oral reading performance.* U.S.O.E. Grant No. OEG-0-9-320375-4269. Washington, D.C.: HEW, 1973.

Goodman, Y. M., & Burke, C. *Reading miscue inventory.* New York: Macmillan, 1972.

Smith, E. B., Goodman, K. S., & Meredith, R. *Language and thinking in school* (2nd ed.). New York: Holt, Rinehart, & Winston, 1976.

Pat Rigg, Ed.D., is at the Center for the Expansion of Language and Thinking, Tucson, Arizona. Elizabeth Taylor is a doctoral candidate at the State University of New York at Albany.

Retellings as Measures of Comprehension: A Perspective

Sharon L. Smith

In his account of how he learned to classify organisms from the European naturalist Agassiz in the middle of the last century, paleontologist Nathaniel Shaler (1907) described the following process. After a thorough interview, Agassiz placed Shaler at a desk alone with a pickled fish and instructions to learn all that he could about it, without reading anything or consulting with anyone. After an hour of turning the specimen this way and that and noting a number of details, Shaler felt satisfied with his performance and, anxious to get on to more serious pursuits, presented himself for examination. But his teacher had no time for him, neither on that day nor during the rest of the week.

At first Shaler resented the neglect, but in time, resigning himself philosophically to his instructor's singularities, he turned back to the fish and examined it further. To his surprise, he began to find out more about it, perhaps a hundred times as much as he had thought possible at the beginning of the task. During the next week, he spent each entire day with his fish and became, so he thought, a master of his subject.

But when, on the seventh day, he got his opportunity to demonstrate all that he knew, his professor merely listened for an hour and said, "that is not right." Suspecting that his teacher was playing a game with him, or perhaps testing his capability for doing hard, continuous work, the student discarded his notes and began the task over, devoting another full week to the careful and systematic examination of the fish. This time he succeeded in satisfying Agassiz. His reward was to receive a half-peck of bones, the parts of some half dozen fish of different species, which he was to fit into their proper order and classify. Two months were devoted to this task, with no comment from Agassiz except for an occasional, "that is not right." Having mastered the fish skeletons, Shaler went on to the task of sorting out some 20 species of preserved specimens, always working on his own with intermittent feedback when he happened to be on the wrong track. It was not until after this task that the student earned the right to read textbooks and discuss his work with his teacher.

At this point, Shaler began to notice an interesting change in himself: "I shall never forget," he writes, "the sense of power in dealing with things which I felt."

Eventually he discovered a flaw in Agassiz's own scheme of classification, an animal that could not be put securely in one category. He reports, "I had a malicious pleasure in exhibiting my 'find' to him, expecting to repay in part the humiliation he had evidently tried to inflict on my conceit." But the master's reply was, simply, "My boy, there are now two of us who know that."

This rather extraordinary account of an educational process, what we might today call a very pure case of individualized learning, may seem to be lacking in a number of features that we take for granted in most classroom instruction. Most notable, perhaps, was the absence of textbooks, at least at the beginning of the student's instruction, and the injunction that he refrain from discussing his work with others. Moreover, he received no lectures, no study guides, no explanations of how to do things right. He was not even told, in any specific way, what he was supposed to do. One might ask, what was Agassiz doing to earn the tuition he was paid for teaching Shaler? Just what was the role of the instructor in this educational process?

The implications of this story, for the narrator at least, seem to be that the student has full responsibility for his own learning and earns the right to have discourse with his instructors and the books they write, by developing his own powers of observation and reason. While this is an extreme view that would never do as standard policy at any level in our educational system, it contains a large enough element of truth to be worth considering. First of all, Agassiz did not teach in the conventional sense of conveying information and then checking to see whether the student had retained it. In some more subtle way, however, it would appear that he did teach Shaler, who went on to become an international scholar in paleontology. While we can see that his teaching approach was *not* to attempt to transfer the contents of his mind to that of his student, it is difficult to characterize exactly what he did do. One could say that Agassiz taught through minimal guidance (setting up the series of tasks), feedback on negative results ("that is not right"), and example (himself, busily concerned with his own work). For the rest, he left it up to the student himself to find out what his mind could do.

It is this fragile principle that seems so important in education, provided that we think of learning in terms of its long range implications. Our technical approach to learning, which has emphasized external control whether by instructor or materials, may inhibit the full constructive participation of the learner in the development of his own thinking processes. While this statement is not intended to be a global attack on today's educational practices, it is a caution to be heeded by those who deal with professionals in education and publishers of educational materials. Think of Agassiz simply saying to his student, "look at your fish," and ask what this approach might add to our knowledge about learning and cognitive development.

It raises the deeper question of what education should do. In this context, the results of a successful education are seen as a gradual strengthening of the ability of the individual to operate within the structures of his own thought processes, making use of whatever information that he has abstracted and assimilated as a source of power in fashioning constructive solutions to the problems he encounters in school and in life. In this process, he begins with his own mental terrain, the mountain he must climb to gain a broader view of the world. To do this, he must appropriate his own meanings and assemble them into his own logically sound belief system. As educators, we try to guide him in the appropriation of meanings and the structuring task of building a coherent body of knowledge. And now the question becomes: How much of the total work involved should the educator do?

This question is complicated by the fact that in any content area there is a body of information and ideas to be taught. This body of knowledge is what constitutes a "discipline," an elusive concept which has, nevertheless, reality in serving the purposes of organization, consistency, and history. It is a form of human cooperation across time and distance, a collective disposition to merge minds into a mainstream that, it is hoped, is heading toward the ever expanding enlightenment of the species. Yet to the developing mind the "disciplines" have, at best, only a secondary reality. They provide a model and a reference system but do not themselves reflect the structures of the individual's own knowledge. It was for this reason, perhaps, that Agassiz withheld secondary sources of information until his student had gained extensive experience by encountering his object of study in his own terms, but then encouraged the student's interaction with discourse.

But this is the point where most of us as educators begin, with the use of verbal sources as vehicles for learning. With notable exceptions, such as science and vocational classes, we have no analogue for the pickled fish or basket of bones. Our major concern is with books and their relatives, and this leads us to another question: Does Shaler's example apply to the learning-from-reading process as well as to direct observation of phenomena? Can we say, "look at your book," and expect anything like Agassiz's results?

The literature suggests that the "psycholinguists" are answering the question with a confident Yes, while the "cognitive psychologists," more in keeping with scholarly tradition, are saying Yes and No. (The quotation marks in the previous sentence are to indicate the folly of putting thinkers into discrete groups.) Actually, both would seem to agree that what book you place in front of a student makes a difference; it must, for example, be written in his own language, and there are a number of features to be considered as well. What these other features might be is the subject of debate and not the focus of the present discussion. The point here is that there is both an individual and a collective side to knowledge development. Each aspect affects and alters the other. Print, the fact that people read, and the further fact that reading is the main course of school learning are artifacts of this central circumstance. What we all want to know is how we can bring readers and texts together in a dynamic way.

The importance of this problem is perhaps why schema theory has had such an exciting effect on researchers in reading. It is generally acknowledged that the basic tenets of schema theory are not new. We express our gratitude to Kant, to Bartlett, to the Gestaltists, and to Piaget for having articulated the ideas years ago. Yet schema theory is not simply a modernized restatement of what wise scholars had been saying all along, for it incorporates the subsequent development of knowledge in information theory, linguistics, semiotics, and semantics.

and links up with work being done in the simulation of human intelligence. These connections promise to help us penetrate the surface level of broad-based theory and go into depths of articulation well beyond the conceptualizations of earlier thinkers. And this is exactly the way that schema theory would predict the process would work. We make these earlier meanings our own, and then go on to create our own knowledge.

A major assumption in both education and epistomology is that original thinking grows out of established thinking and bears a relationship to it by confirming, modifying, or rejecting it. The traditional "review of the literature" in empirical research is an appropriate example. This assumption gives us a point of view in looking at any school reading, or reading in order to learn. What we mean by learning at elementary and secondary levels is really the same kind of thing that we mean by discovery of knowledge at the scholarly level. In both instances, reading is a means to an end. From this it can be inferred that, to understand the reading process, we must examine the linkage between the source of given information and the response that the reader makes to it. This linkage is not the same as overlap. We want to know as much as we can about the text to which the reader is exposed, and as much as we can about his own operations upon the material; but simply trying to match the two does not really get at what we want to know about the development of cognitive skill or cognitive power that might occur as a result of the operational process we call reading.

When two people are having a conversation, it is not usual for one to repeat what the other says in order to establish comprehension. Rather, the two conversationalists respond to each other, and the event is, in Rosenblatt's (1969) term, transactional. Sometimes one may say something along these lines: "While you were talking, I thought of this," and then go on to explain his idea. Such a way of dealing with incoming information is a natural way of demonstrating comprehension. If the two speakers are not hearing or understanding each other, this will become apparent in the lack of connection between their contributions to the discussion. Is not a similar transaction characteristic of reading comprehension?

This observation might raise the question of whether we are, at this point, talking about reading at all, or rather some other cognitive phenomenon that happens as a result of it. This question, it seems, is valid but not perplexing. First of all, we are concerned about the operations that occur during reading, although these cannot be directly observed and must be assessed after the fact. But we are also concerned about the comprehension and integration that occur after reading, and the way that reading activates existing knowledge and causes rumblings, uprisings, and resettlings in the mind. Reading cannot be isolated from these other cognitive events if for no other reason than that it is done specifically to cause them.

To return to schema theory, which does connect reading comprehension with prior and subsequent acts of comprehension, there are two conceptualizations that are useful in thinking about school-related reading: top-down/bottom-up processing, and high/low text integrity. The first, which suggests that sometimes readers build up structures of meaning piecemeal and sometimes begin with structures already at hand, gets us out of certain either/or debates that have long since served their term of usefulness as clarifiers of issues.

Schema and other reconstructive theories of learning also suggest that long-term meaningful learning depends on the integration of new information to existing knowledge structures. Those who support this theory also point out that certain kinds of memory tasks lead to the isolation of material from other knowledge in an effort to preserve its integrity and avoid errors in recall. However, this isolation or "freezing" of new information, by keeping it out of relationship to the individual's operative network of meanings, keeps it from becoming functional in his need to update his personal knowledge system.

Learning, according to schema theory, consists of updating knowledge—that is, modifying existing schemata and creating new ones (Anderson, 1977). This process operates most smoothly when the learner is able to subordinate incoming information to existing knowledge structures and allow interaction to take place. However, when the integrity of the text takes precedence over the conceptualizing activity of the learner, and he is directed to maintain the input in as unaltered a form as possible, he will be able to do so only by isolating this processing task from his normal intellectual functioning. This freezing of the normal cognitive processes is likely to occur in memory experiments and in some classroom learning tasks, where the concern is with accuracy of text recall (Anderson, 1977; Spiro, 1975). According to schema theory, such information would have little functional importance for the learner in his concern for updating his own knowledge and so would result in little long-term learning.

Schema theory suggests that it is the natural knowledge-building activities of the learner that contribute to the growth and development of the mind. If so, classroom instruction should center on the conceptualizing behavior of the student. Reading for recall, to be sure, still has an important place in classroom learning. Since the main interest of the learner, however, is to incorporate information into his own developing schemata, activities that enable him to bring these schemata into operation during the processing of information will have the most permanent value for him.

Consider, at this point, the following statement written by a college freshman describing his difficulties with chemistry:

> One of the reasons why chemistry is so difficult is that it is so abstract. Chemistry is a topic discussing something that we can't see or smell or feel. It is a science based on theory that can't be proven or disproven [by the individual learner]. It is something that you just can't relate to. . . .
>
> The reading in this area is often boring, with much information in each line, usually difficult to comprehend. Chemistry, as with most other sciences, requires recalling previously learned information, which is applied throughout the duration of the course.
>
> The only way to deal with the reading is to study it and spend more time with it. Try to stay up with it and remember what you learned in the past. This should help the problem of recalling and applying previously learned material.

> One of the ways to remedy the problem of it being so abstract is to look at it in terms of something. For example look at molecules as eggs, something we can feel and touch and are able to work with. My chemistry teacher once told me that a mole of people would cover the face of the earth fifty miles thick, and that a mole of atoms could fill a glass jar. This makes it easier. Chemistry is chemistry and it will always be boring. But what you have to do is try to think of ways to apply it to everyday life. For instance, why does salt melt ice? Looking at it in these terms can often make it more interesting.

This student began the class with no apparent schema for its content, so he takes a bricklayer's approach, trying to build his structure from the ground. But when he tries to connect the new information with structures of meaning that already exist in his experience, the construction task becomes easier (though still far from exciting).

Following is another account by a student experiencing similar problems of involvement in a computer science course:

> The major problem I have is that the material is all new. The class moves very rapidly and I often feel unfamiliar with the terms and concepts that are necessary to continue. The Fortran language itself is like an entirely new language, more difficult than mathematical equations, and to function in the class I must be able to use it fluently. Finally assignments are extremely long and difficult.

This student describes his solutions to these problems as follows:

> Material in any class is bound to involve new and challenging material. I can approach the class as a fresh challenge. By using the material at every opportunity, I may be able to familiarize myself with it more quickly. Also, by finding applications relative to myself, I can create a strong desire to learn the material. I obviously must keep up with the quick pace. By determinedly sticking to it, I'm sure I'll be able to keep up. Also I can take advantage of my free time to get ahead. If I've created a strong desire, this might be a treat. The most important strategy would be to not allow myself to fall behind. As for using the Fortran language fluently, I must rub my nose in it and plow through it over and over till I master it. When working on assignments I can speak and think in Fortran. The assignments are long and are bound to take time. I must allow for this when making out my weekly schedule. Perhaps I can work with classmates, which would make it more interesting and economize my time.

While one might detect a certain amount of bravado in this student's plan, he is doing his best to evoke some kind of schema for the situation if not for the content itself. His schema here is an image of himself as an individual who gladly accepts challenges and enjoys overcoming his own fallibilities. (Shaler did this too in his early encounters with the fish.) If bottom-up processing is to be his lot, at least until he achieves fluency in the artificial language, he will make the best of it. It probably helps, though the student doesn't say so, that knowing this particular subject matter will later prove useful in finding employment. The chemistry student probably could not make so direct a connection. But beyond this kind of motivation, there is a sense of experience in the second writer's account. He has been through it before, and by sheer diligence has probably constructed schemata for other subjects that required nose-rubbing and plowing at the beginning. This gets us into an issue of considerable interest in education, the problem of getting students started on schemata-building for relatively new areas of information. Where does the energy come from, or the willingness to mobilize existing energies? How does the learner get started, and what guidance, if any, will help?

One way to begin this line of inquiry is to look at what students do when they encounter course content, for example, when reading textbooks. Since we cannot actually view the workings of the mind during reading, we must rely on some product or result from reading as evidence of the process. Open response verbalizations (retellings, oral or written) will provide the "purest" form of reader response in the sense that no information or language has been contributed by a test question as would be the case with any type of conventional test item. This can also be a problem with probed retellings, in which the probes might provide "stems" to which students connect their responses. The reason why this might be a problem is that any such direction on the part of question or examiner will tend to specify the type of response given (frequently recall), so that the verbalization is no longer necessarily representative of the operations performed by the reader as he attempts to assimilate the information. On the other hand, in an unprobed retelling a student may not reveal all the information he has, especially if the mode (speaking into a tape recorder or writing) is uncomfortable for him. However, if the objective of the assessment is to find out *how* the student is thinking about the content rather than *how much* he can demonstrate that he knows, the unprobed retelling is still the best (in the sense of least contaminated) response. That is, the unprobed retelling will reflect the reader's independent interaction with the text without interference from a third party.

The analysis of the retelling, then, should not emphasize recall in the form of repetitions or paraphrases, although these will be of interest in the analysis. The focus here is on the reader as author of his own version of the content, which will indicate the nature of his own active structuring processes. The analysis, then, should attempt to reveal how the reader is breaking meaning out of the text's language and then reconstructing it in language of his own. This is important if we accept the principle of top-down, as well as bottom-up, processing in schema theory.

In pursuit of this idea, the writer conducted two studies designed to look at what it is the reader does in his efforts to verbally conceptualize text content, one with college seniors and graduate students and one with high school sophomores (Smith, 1977, 1979). Because of the considerable differences in age, academic experience, and type of text used, the outcomes of the study also differed greatly. The older subjects read a politically oriented magazine article that required the reader to provide his own generalizations and tended to elicit a strong personal response either favorable or unfavorable to the author's point of view. In their unprobed retellings, these subjects produced discussions of the content (as opposed to paraphrases of it) that were relatively long and contained few propositions from the text itself. To the extent that these results can be considered representative of the cognitive behaviors of "good readers" at a mature level, this study seemed to support the concept of reader as author of his own text.

The second study involved much younger readers, who were responding to a passage from a supplementary text in American history. The passage was chosen on the basis of a number of criteria: relevance to subjects' current course content, readability (Fry, 1978) at a level slightly below 10th grade, variety of types of format (prose, tabular, and pictorial) and primarily factual content, with little interpretation given in the text. As in the earlier study with older college students, unprobed retellings were analyzed to determine the number of recall, inferential, and supplemental responses. With this group, the retellings were shorter, more text-bound, and apparently not connected with other things that students knew or thought about. Written retellings done 10 days after the passage had been read showed that certain kinds of global evaluation responses (e.g., "I didn't like it," or "I thought it was interesting") tended to replace recall responses, but there was still no convincing evidence, other than amount of recall after a certain lapse of time, that students had assimilated the text into their own knowledge structures. This observation led to some tentative comments on the learning-to-read process in school. Also, the data collected, oral and written retelling protocols, provided some insights into students use of their own language for learning.

When a reader discusses the content of a text he has read, he is fitting his language to the information, trying to find a match. In doing this, he has access to a great deal more than just the information given in the text, or the version of it that has been created in memory as a result of reading. There is perhaps no good way of sorting out specifically what information has been deposited as a result of a specific reading experience and what information comes from other sources. But this distinction probably does not matter a great deal when the focus is on learning as it evolves over time. When a person is in the position of speaking or writing about what he knows, without a specific text in recent memory that he feels he must represent in as complete a fashion as possible, the source of the information is not the issue. Rather, it is the construction that he is now capable of producing. This approach suggests several questions.

What is the role of language in conceptualizing? What is it that the reader has in his knowledge of his native language that enables him to have a representation of information available for communicating it to others? What is lost (and gained) as information is stored and modified in long-term memory? In what way does mode of communication affect the role of language in learning? In what sense is a reconstruction a learning experience? What non-cognitive factors (assuming that it makes sense to distinguish cognitive and non-cognitive factors, when speaking of the operations of the human mind) are involved in learning? And what are the implications of these questions and their answers for instruction?

Looking at unprobed retellings is an attempt to focus in on the junction of the communication process—the point at which what has been acquired or stored is put into the form that can be relayed to another person. It might be called the "language relay." When a person communicates verbally, he is searching throughout the process for both ideas and for language to express these ideas. It is a complex process and often a difficult one. In trying to express an idea, the person often reveals his search for words, and the search for words can itself interfere with thinking. Perhaps it is comparable to stumbling over crucial vocabulary items in discourse reading. The following oral retelling of a text about recruitment procedures in the Civil War reveals the speaker's search for language at certain points in the reconstruction, which represents a good attempt to convey the main ideas of the passage.

> Okay, this reading was about the different reactions that people had about the draft during the Civil War. And, at the first they didn't have any problems because the Union was winning, and they didn't have any trouble—they didn't have to draft anybody. But then, they lost an important battle, trying to capture the Confederate capital of Richmond, and they had to start drafting people. So, in 1862, Lincoln asked for 300,000 men to go into the draft, I think it was, and he only—90,000 responded. So they had to start drafting people, and, they could—you could—get out of it if you could find a substitute or you could pay a discharge of $300,000. But most of the people didn't have that kind of money, 'cause they were immigrants and they were poor and couldn't pay it. And, then, people started getting excuses and making, I think it was, exemptions and stuff, and he had to make a—Lincoln had to make a proclamation saying that, about all the different things you couldn't do, that were illegal about the drafts, and you didn't do them, then you were subject to martial law. And they also had a lot of problems in New York where the first draftees were written in the newspaper, and the mobs attacked the newspaper place, and they killed some blacks because they thought they were responsible, since the war was fighting over slavery, and a whole bunch of other people were killed. And the rest of the tape [passage] goes on to tell about, just statistics of the different—how many draftees there were, how many substitutes and—then it talked about the Confederate, um-m, part of the draft thing. And their program was just about the same, except, I think, they could have one exemption—one overseer or something like that. And then it tells about how some of the men felt about the draftees—men that were volunteers—how they felt about the draftees, 'cause they didn't like it, 'cause they didn't like working with people who didn't want to support their country.

One can easily sense in this retelling where the discussion is smooth sailing, where the speaker found the going rough, and how the speaker recovers himself at these points. Also, it is evident in this reconstruction that the speaker is alert to causal relationships in the information and uses these as a logical basis for his retelling. Even without a formal analysis, one can see that coherence is maintained largely by use of conventional linguistic/logical cues such as conjunction and subordination.

Consider in contrast, what is happening in another reader's written reconstruction of the same passage:

> The North recuted 300,000 men, but only 90,000 showed up. And thay would go around buring down houses and burning affings [orphanages]. And the 2nd time that recuted they got a bigger resoneds. To me I thot it was really bad all of them going around burning up all those buildings and affunige. Thay killed thousands of people men women. Each farmer was allowed 15 to 16 slaves each.
>
> To me all together it seems really dumb for all thoes peopl to be fighting like that cann't thay just shar the land tay could have diviede it and thave there wont had been any fighting and I think it was alful that thay treted all thoes

> people like that not just the white but the black to the had them as slaves and beated them and I think it was really mean.

Spelling is only the most visible problem here. What is most seriously wrong is an inability to fit written language to the information that the reader has picked up from the text. Also there is no order in the presentation of different items of information in the first paragraph, which is the writer's attempt to represent the information he recalls. He has not retained an operational version of the logical relationships that give meaning to the factual content. In this passage it seems apparent that the difficulties experienced by the student in trying to use written language as a form of communication interfered almost completely with the knowledge he had of the text.

As these examples suggest, looking at (or listening to) a reader's verbalization of text content can lead to useful insights concerning his language processing behaviors as these relate to the assimilation of information and concepts from reading. These observations can be related to a conceptualization of the learning-from-reading process in which the text is represented by a center bounded by "regions" of inference and supplementation. Learning proceeds in a wave-like fashion, going above and below a line more often than staying directly on it.

Following are the beginnings of 10 oral retellings of the Civil War passage chosen at random from 48 protocols by students who did the retelling immediately after reading the text. Each excerpt given here consists of the first 100 words of the protocol, plus the rest of the last sentence. Students had been instructed to retell the passage to a person who had not read it.

The reason the beginning of each passage was chosen for examination here is that this part represents the student's approach to construction of his own text. What we are looking at here, in effect, are these students' attempts to make use of the information gained in one text in the construction of another. The original text is not given, so comparison of the retellings with the source passage is not possible here. The only comparisons that can be made are among the excerpts themselves. The question now is this: given simply the productions of students' retellings after reading, what observations and questions can we derive?

Example 1

> Okay, they started the draft in 1862. The first one was July 1, 1862. Before then Lincoln had asked for 300,000 volunteers for a three year enlistment. But he only got—he got a bad response, so he made it down to just nine months, and he only got 90,000 people. So he called for the draft for 300,000 people, and he called to the governors to give him the people. They didn't want to draft or anything. They thought that if they had a draft they would lose the war if they had to draft men. Let's see. South drafted men too, but the South had a different point of view.

Example 2

At first, in the beginning of the war, the Union had a little trouble drafting—little trouble finding volunteers. Grant was winning in Virginia, I believe it was, and everything was going fine. But then, Secretary of War Stanton made the assumption that the Union was winning and thereby needed no more volunteers for the duration. And the Army of the Potomac, I believe it was called, suffered a defeat. Both of these things tended to reduce the number of volunteers. So Congress had to pass a Conscription Act in which it asked—in which each state was to fill a certain quota either by volunteer or draft or a combination of the two.

Example 3

It described what the draft laws were like in the Civil War during 1863. President Lincoln asked for so many men to be enlisted for a period of three years, and this was the number of men he wanted: approximately 300,000. Then, when he got no response to that, he reduced the time of enlistment to approximately nine months. And it describes how Grant had won battles at two forts on the Tennessee River and the Cumberland River, and how the Army of the Potomac had failed in its mission to capture the Confederate capital at Richmond.

Example 4

This is a story about the new draft law that they made when Lincoln was President. And before they had it what they did was, they asked for volunteers, and Lincoln asked for 600,000 volunteers one year and nobody really wanted to go, so they had to make up this new law, I think it was called the "Conspiration Act." That was where the people's name that were picked, they had to go to war, they were drafted to go to war. And the only way they could get out of it this was if they could get a substitute to go for them.

Example 5

Okay, Lincoln, he wanted 300,000 men to sign up for instead of six months he had said earlier, for three years. And like, not very many men hired at all. And so he changed it from instead of three years to nine months, and only 90,000 came. And so he had to get a Conscription Act that said that if you were disabled or, anything like that, you didn't have to sign up, you didn't have to be drafted, and you could also buy someone else in your place for $300.00. And a lot of people did. And um-m it started getting better because more people would join in because they were paid for, someone paid for them, or they got benefits, or something like that.

Example 6

It's about, like Lincoln,—it was about President Grant, first he captured two important cities and he, like, they were going to get drafted if they were like between the ages of 20 and 45 (the men). And if they didn't want to be drafted they either had to pay $300.00 or else bring in a substitute. And then, okay, so when they were going to get drafted, I guess, some of the people that had to get drafted they want to, they burned down the building of the people that drafted them. And so then that was the blacks that did that, and they were against blacks.

Example 7

Starts off that the Civil War is going on. And at the beginning of the Civil War, there were plenty of recruits for the war. But as times goes on and they start to lose a few battles, it looks like it's going to be a downfall, so nobody wants to sign up anymore. So President Lincoln asks again for 3,000 men to go a one year term—a three year term I think it is—in the army, but nobody wants to do it. So then he asks again for 3,000 men to go a nine year term but less than 90,000 sign up.

Example 8

This paper, I guess, was about the ways in which the North and the South attracted men to enlist in the army. Some of it was drafting, and first was voluntary, until the North decided that they needed more men because they—because their draft—I mean their voluntary—volunteers decreased in amount. And so they advocated a draft, like 300,000 men from the states for the North. You could either find a substitute to go in your place if you didn't want to be in the army or you could pay a $300.00 fine. Let's see, South, they had a draft but you could only find a replacement, you couldn't pay a fine.

Example 9

The draft—okay—Lincoln wanted 300,000 men to serve, and so nobody showed up. It was a bad response. So in August—I'm not sure what year—he asked again for 300,000 men, but he only wanted them to serve nine months. And there were only 90,000 I think showed up. And so the governor decided that they couldn't have it. So they made the draft, and that was in—I don't know when that was. Okay, and the people didn't like that too well. And they picked people from 20 to 45.

Example 10

Okay, this reading was about the different reactions that people had about the draft during the Civil War. And at first they didn't have any problems because the Union was winning and they didn't have any trouble—they didn't need to draft anybody. But then, they lost an important battle, trying to capture the Confederate captial of Richmond, and they had to start drafting people. So in 1862, Lincoln asked for 300,000 men to go into the draft, I think it was, and only 90,000 responded. So they had an act that said that you know, they had to start drafting people.

Considering reader retellings as texts immediately suggests the basic question we ask of any text: Is this person saying something that I (the listener/reader) can make sense of? Judgment of this issue is likely to be made on at least three criteria: (a) sufficiency of the context within which specific information is presented, (b) text cohesiveness, as discussed by Halliday and Hasan (1976), and (c) text completeness, an effect, possibly, of the relationship between new information and redundancy.

We can define context as the framework or structure that connects specific information to other knowledge. Context is constructed both within and beyond the language of the text and requires the cooperation of both text producer and text comprehender. In Clark's (1975) words, it is a product of the "given-new" contract. Looking at just the first sentence in each of the excerpts above, we can infer how the speaker has perceived his responsibility for the construction of an explanatory context. For example, Speaker 1 leaves it to the listener/reader to provide the reference for "they" and to identify the situation that called for a draft. For accommodation, Speakers 8 and 10, perhaps, do the best job. These statements attempt to orient the listener to the broader situation within which this discussion of drafting occurs and establish a point of view.

Discussing the concept of cohesion, Halliday and Hasan (1976) define a text as "any passage, spoken or written, of whatever length, that does form a unified whole" (p. 1). While they admit to various degrees of "textness," they believe that on the whole we can decide with reasonable agreement whether a given passage constitutes a text. Since unity is a basic premise of their concept, they argue that cohesive "ties," e.g., pronomial reference, repetition, and other linguistic elements that assume their meanings from other parts of the text, provide the weavery that pulls a text together. Sequence and logical relationships among statements also are important elements of cohesiveness, and context or framework plays an obvious role.

The third criterion, completeness, refers to the adequacy of the content itself. It is described above in terms of a relationship between information and redundancy to indicate that the person toward whom the text is directed seeks a comfortable mix of new and previously stated information. In this aspect of the reader/author's production, we can gain some sense of his operations on the content as these affect his storage and retrieval processes.

Whether these or other criteria are used to construct a framework for looking at reader productions as texts in their own right, the focus should be on how (not what) the reader comprehends. This focus is possible only if we assume that comprehension is not an isolatable phenomenon but is always related to expression. Language itself does not drop into the mind. Rather the reader cracks the meaning out of the language of the text and puts it into language of his own.

Rosenblatt (1969), in her exposition on the reader's response to poetry, speaks of "a reading that results in a work of art." In speaking of a reader's response to expository prose, can we not also speak of a reading that results in an informational text?

Rosenblatt's thesis applies very well in both cases:

> The coming together of a particular text and a particular reader creates the possibility of a unique process, a unique work. (p. 43)

The way that cognitivists are looking at reading is comparable in certain respects to field theory in physics, which views the operations of the universe in terms of a field of activated energy rather than a mechanical model with discrete components. Theorists in reading, too, are disposing of component-based models of the reading process in favor of the view that at any level (e.g., sound, letter, or word) apparently discrete units lose their individual character in the flow of meaning, much in the same way that atomic particles lose their apparent solidity and assume wave-like configurations operating under laws of probability. In other words, the relativistic view of the universe is applicable to the nature of meaning in human communications. The notion of absolute meaning as contained in the text is as inadequate as the notion of absolute matter, space, and time. The idea that meaning in the text can "cause" the existence of meaning in the mind has no greater validity than the Newtonian concept of a strictly causal basis of natural phenomena. It might even be useful to think of the human mind as curved space, where classical geometries are no longer operational.

At the turn of the century, William James also cautioned against evaluation that centered on measures of recall, and even went so far as to praise the mind that fumbles with remembered detail.

> We are all too apt to measure the gains of our pupils by their proficiency in directly reproducing in a recitation or an examination such matters as they may have learned, and inarticulate power in them is something of which we always underestimate the value. The boy who tells us, "I know the answer, but I can't say what it is," we treat as practically identical with him who knows absolutely nothing about the answer at all. But this is a great mistake. It is but a small part of our experience in life that we are ever able articulately to recall. And yet the whole of it has had its influence in shaping our character and defining our tendencies to judge and act. Although the ready memory is a great blessing to its possessor, the vaguer memory of a subject, of having once had to do with it, of its neighborhood, and of where we may to recover it again, constitutes in most men and women the chief fruit of their education. This is true even in professional education. The doctor, the lawyer, are seldom able to decide upon a case off-hand. They differ from other men only through the fact they know how to get at the materials for decision in five minutes or half an hour, whereas the layman is unable to get at the materials at all, not knowing in what books and indexes to look or not understanding the technical terms.
>
> Be patient, then, and sympathetic with the type of mind that cuts a poor figure in examinations. It may, in the long examination which life sets us, come out in the end in better shape than the glib and ready producer, its passions being deeper, its purposes more worthy, its combining power less commonplace, and its total mental output consequently more important. (James, 1899, pp. 69-70)

Learning involves change, not reproduction. The interaction of two (or more) minds via a text is dynamic process for the reader, who takes up the act of knowing where the writer left off. If we are interested only in catching photographs of the original text through reader recall, we are confining ourselves to a static view of comprehension that tells us more about our own observational procedures than about the mental experience of the reader. To capture something of the nature of the reader's self-guided interactions with print, we must resist controlling the outcome as much as possible, either by insisting on recall or directing responses through probing. Although it is true that the investigator is always a participator in the phenomenon being observed, this role can be modified as can any other aspect of the investigative procedure. Since the reader and investigator are in a mutual enterprise, sharing the field so to speak, the investigator can, by controlling his own participation, expand or restrict the free functioning of the subject's activity.

A major problem with textbook study is that somehow it is experienced by many students in a static way rather than as a dynamic process that produces change. Let us go a step further and say that information itself is a process, interacting with other ongoing processes in the mind of the reader. This view enables us to look, always, for originality and relationships at the same time. The response that represents comprehension is thus always new but never isolated from the information processes with which the response is involved. The value in focusing on the response processes lies in their character as representative of the learner's own mental experience. We can dissect these processes in an effort to identify which ingredients came from the external stimulus (e.g., the text) and which from the inside, and this approach often has pedagogical value. But in so doing we must be aware of the distortion introduced by stopping the processes, as it were, in order to examine the parts as a method of explaining the processes. As an alternative, there are times when we should simply keep the processes as intact as possible and see what we can learn from them.

References

Anderson, R. C. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Clark, H. *Comprehension as an art*. Indiana University psychology forum lecture, Bloomington, February 28, 1975.

Fry, E. A readability formula that saves time. *Journal of Reading II*, April 1978, 513-516.

Halliday, M. A. K., & Hasan, R. *Cohesion in English*. London: Longman, 1976.

James, W. *Talks to teachers on psychology*. New York: Dover, 1962. (Originally published by Holt, 1899.)

Rosenblatt, L. M. Towards a transactional theory of reading. *Journal of Reading Behavior*, 1969, 31-51.

Shaler, N. S. *The autobiography of Nathaniel Southgate Shaler*. Boston: Hougton Mifflin, 1907.

Smith, S. L. *An analysis of reader operations in the processing of extended prose*. Indiana University, unpublished manuscript, 1976.

Smith, S. L. *Learning from reading at the secondary level: An investigation into cognitive operations employed in the reconstruction of text content*. Unpublished doctoral dissertation, Indiana University, 1979.

Spiro, R. J. *Inferential reconstruction in memory for connected discourse* (Tech. Rep. No. 2). Urbana-Champaign, Ill.: Laboratory for Cognitive Studies in Education, Center for the Study of Reading, University of Illinois, October, 1975.

Sharon L. Smith, Ed.D., is director of the Reading and Study Skills Center and assistant professor of education, Indiana University, Bloomington.

Comprehension and Predictability: An Analysis of Beginning Reading Materials

Lynn K. Rhodes

The study presented in this article investigates the effect of the use of alternative theoretical models of reading, which has been shown to pervasively influence reading and reading instruction (Harste & Burke, 1977, 1978; Harste, in press; DeFord, 1978; Mitchell, 1978), on reader/text interaction, more broadly described by Harste and Carey (this volume, beginning p. 4) as the transaction which occurs in the environment of mental setting and print setting. The study was based on several major assumptions:

1. Readers and texts (or authors) have identifiable theoretical models of reading.
2. The author's theoretical model of reading instruction (his purpose in writing the text) influences the semantic structure of the text.
3. The reader's theoretical model of reading influences the strategies applied to the text in the process of reading.
4. The strategies a reader chooses to apply in the reading process will be affected by the semantic structure of the text.
5. Reader comprehension will be affected by the semantic structure of the text.

The diagram in Figure 1 (adapted from Harste, in press) incorporates the above assumptions. As one aspect of mental setting, the reader's perceptions of reading contribute to the selection of reading strategies that reflect the perceptions. The selected strategies are used to sample from the available cues in the structure of the text, an aspect of print setting. In turn, the structure of the text, reflecting the author's purpose in writing it, contributes to a modification of the reader's perceptions of reading in that particular print setting. Those modified perceptions, in turn, modify the strategies selected for use in the reader's processing of the text.

The present study investigated the encounters of beginning readers with texts reflecting alternate theoretical models of reading. Because the children in the study were selected from classrooms employing the same basal reading series, it was expected that their reading models would be homogenous. On the other hand, the texts used for the study were selected to represent three alternate theoretical orientations: phonics, skills, and whole language models of reading.

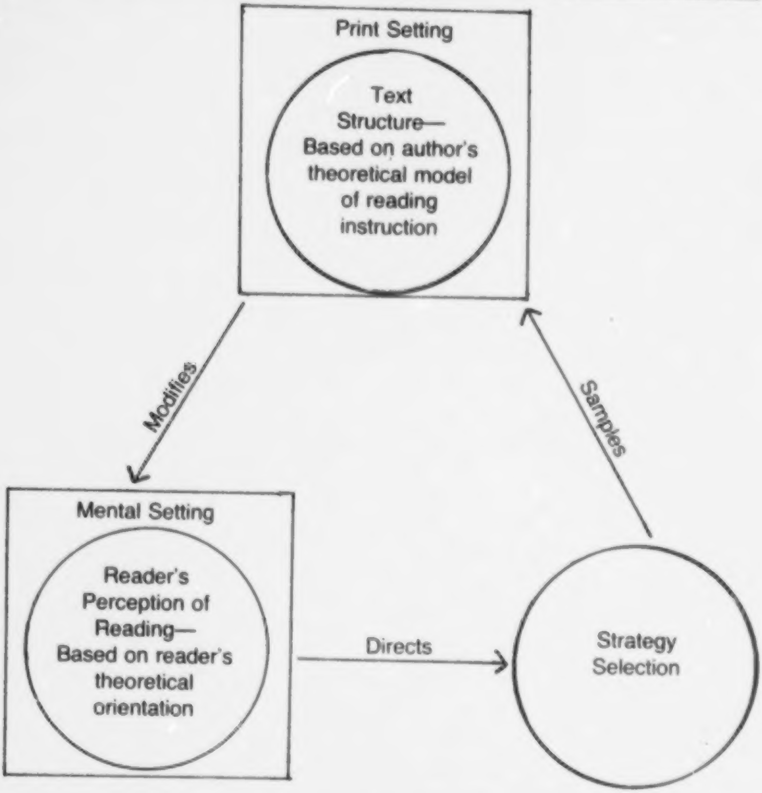

Print Setting
Text Structure— Based on author's theoretical model of reading instruction
Modifies
Samples
Mental Setting
Reader's Perception of Reading— Based on reader's theoretical orientation
Directs
Strategy Selection

Figure 1. A reader encounters a text.

It was hypothesized that the beginning readers would more effectively read texts reflecting whole language theoretical orientations (more predictable stories) than texts reflecting skills and phonics orientations (less predictable stories). Because an understanding of the three theoretical orientations is vital to an understanding of the results of the study, the models are outlined prior to a presentation of the study itself. The exposition of each model includes a description of the text(s) selected in this study to represent the particular model as well as descriptions of reader's perceptions representative of each of the orientations.

Theoretical Models of Reading

Theoretical orientation, as defined here, is the language user's perspective concerning the primary function and focus of written language. The three models used in this study, as postulated by Harste and Burke (1977), can be placed on a continuum (see Figure 2) and are identified primarily by the language focus and unit emphasized by proponents of the model.

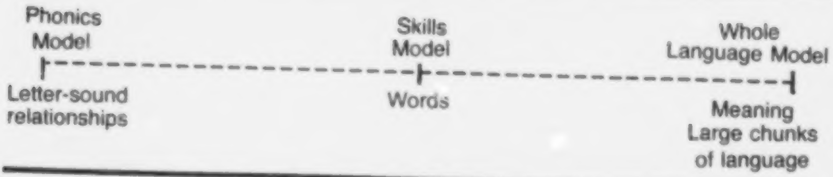

Phonics Model
Letter-sound relationships
Skills Model
Words
Whole Language Model
Meaning
Large chunks of language

Figure 2. A continuum of theoretical models of reading.

The Phonics Model of Reading Instruction

The focus of the phonics model (Figure 3) is on letter-sound relationships, the graphophonic system of language. Association of letter-sound relationships is the initial emphasis in instruction and prerequisite to word and meaning identification and instruction. Phonics proponents believe that language can be broken down into parts to be examined and learned by children; relationships can be learned in isolation from meaning and even words. It is believed that language is learned by building from the smallest to the largest unit.

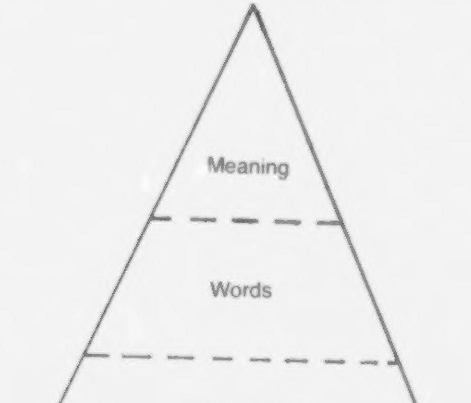

Meaning
Words
Letter-sound relationships

Figure 3. The phonics model.

Reader strategies reflecting a phonics model. In the current study children's reading models were identified by means of *The Reading Interview* (Burke, 1978). The ambiguous term "something" in the first question of *The Reading*

Interview compels students to identify the unit of language being focused on as reading strategies are verbalized.

An example of one subject's answer that focused on letter-sound relationships is:

> Interviewer: When you are reading and you come to something you don't know, what do you do?
>
> Shane: Try to sound it out.
>
> Interviewer: Do you ever do anything else?
>
> Shane: See if the vowels have the "e" marker on it. If the vowel is long, the "e" would be on the end of it.
>
> Interviewer: Anything else?
>
> Shane: Try to sound all the letters you can.

Information from the remaining nine questions of Shane's interview confirms that he holds a phonics model of reading. For example, he believes that the best reader he knows, his teacher, sounds out words when she comes to something she doesn't know. And, in order to help a friend who is having difficulty reading, Shane would assist his friend in sounding out the words.

Texts reflecting a phonics model. Phonics basal reader programs feature the letter-sound relationship as the primary focus of instruction, the basis of all other reading instruction. The texts or stories that are a part of such a reading program are written with carefully controlled vocabularies containing previously introduced letter-sound relationships. The authors of phonics texts believe that supportive, readable texts are those which contain letter sequences that children have learned to recode (sound out).

Because some of the most frequently used words in the English language do not follow regular letter-sound correspondence rules, some words are introduced in phonics programs as sight words. These words, once introduced and practiced, are entitled to become part of future texts.

The text selected in this study to represent the phonics model was "The Big Stand," a story in *Keys to Reading* (Harris, Creekmore, & Greenman, 1972a), published by Economy Co. The *Keys to Reading* teachers' manual notes that "the introduction of sounds and spellings is carefully controlled" (p. 6) and that such a phonics model is preferable to a skills model where "attention to the critical discrimination skills is delayed until a sizable group of sight words has been memorized" (p. 5).

The vocabulary used in the reader containing "The Big Stand" includes "reading vocabulary" (sight words) and "sounding vocabulary" (words with a regular and previously introduced letter-sound correspondence) in approximately a one to three ratio.

The story itself, "The Big Stand," reflects the constraints, minimal content, and stilted sentence structures that such a phonics orientation places on a text.

> Don came to the stand.
> "I want something," said Don.
> "I want something to eat.
> What can I get from this stand?" (p. 30)

The content of the story suffers; the tightly controlled vocabulary renders the story vague and somewhat incomprehensible. The language is stilted because the authors are operating on the belief that short sentences constructed of a string of phonetically regular words are easiest for a beginning reader to cope with.

The Skills Model of Reading Instruction

The second alternate orientation to reading instruction is represented by the skills model (Figure 4). The primary language unit in this model is the word. It is believed by proponents of this model that meaning automatically follows recognition of a string of words in a sentence; that meaning is, in fact, the sum of the meaning of each of the words in a sentence. As in the phonics model, it is believed that word recognition is a necessary mediation to meaning. Because reading is viewed as the sum of learned skills, word recognition, grammar, and comprehension, skill hierarchies have been developed by proponents of this model to facilitate learning.

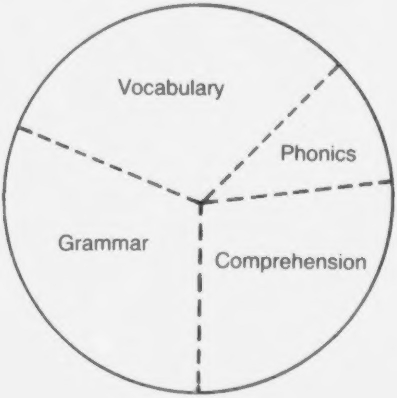


Figure 4. The skills model.

Reader strategies reflecting a skills model. In *The Reading Interview*, (Burke, 1978) readers who hold skills models identify the word as the primary language unit to be focused on in the process of reading. Shanna, one of the children in this study, held a model that was closest to the skills model point on the continuum.

> Interviewer: When you are reading and you come to something you don't know, what do you do?
>
> Shanna: Ask my mama what the word is.
>
> Interviewer: Do you ever do anything else?

Shanna: I figure it out.

Interviewer: How do you figure it out?

Shanna: Spell it. Like if the word was my name, I'd say S-H-A-N-N-A and then I'd know what it was.

Shanna's preferred strategy when she comes to a word she does not know is to ask someone what it is. Though she also mentions a sounding out strategy (albeit an unusual one), answers to other questions in her interview make it clear that asking for unknown words (and telling others words) is her preferred strategy.

Shanna has the notion that words are either known or unknown. In fact, she later states that her teacher, the best reader she knows, never comes to an unknown word in the process of reading. For Shanna, reading instruction consists of learning to recognize some finite list of words.

Texts reflecting a skills model. The primary language unit in skills basal reader programs is the word. Texts or stories are constructed primarily of known words, that is, words that have been introduced in previous lessons and practiced repeatedly. Prior to each story, teachers introduce new vocabulary, words to be added to the child's sight word vocabulary. Because practice is considered the key to immediate word recognition and immediate word recognition the key to comprehension, words are repeated again and again within stories and from story to story.

"Pork Chops and Applesauce," a story appearing in Ginn's Reading 360 (Clymer & Jones, 1973a) basal reader *Seven is Magic*, was selected for this study as representative of the skills model. Stories in this program are constructed from a controlled vocabulary list consisting of basic (sight) words, decodable words, and enrichment words. Once new words are introduced, they are "given reasonable repetition" (*Teacher's Edition: Seven is Magic*, p. 24) to ensure recognition in future stories. This is not true of enrichment words because they are "of low utility for the rest of the program" (p. 24). The learning of sight words, traditionally the basic language unit of skills programs, is supplemented by decoding skills "introduced systematically" because "recent research and curricular developments require that the older method of word presentation be modified by an approach which reveals phonemic and structural aspects of language at an earlier level" (p. 24). The ratio of basic (sight) words to decodable words in the Ginn series is approximately three to one, the opposite of the focus found in the Economy series.

Though the syntactic structures of the story selected for this study from Ginn's program, "Pork Chops and Applesauce," are more natural than "The Big Stand," the text still seems somewhat stilted and vague due to vocabulary control. For example, the reader does not know that the "store" is a grocery until Mr. King gives Ted instructions about where to put the food: "here with the others" rather than "in the meat case."

An excerpt from "Pork Chops and Applesauce" provides a sample of the syntactic structures, the content, and the word repetition:

> Then Ted remembered something.
> He liked applesauce with pork chops.
> Sometimes Mother forgot to get the applesauce.
> Ted thought, "Maybe other people forget it, too.
> I can get some cans of applesauce and put them with the pork chops.
> Then other mothers will not forget." (pp. 36-37)

In short, the skills text is a direct reflection of the traditional notion of readability. The authors of skills texts believe that supportive, readable texts are those that consist of known words (due to repeated previous practice) arranged in sentences that are kept short to reduce grammatical complexity.

The Whole Language Model of Reading Instruction

The third alternate orientation to reading instruction is represented by the whole language model (Figure 5). The core of the whole language model is meaning, the semantic system. The syntactic and graphophonic systems give predictable form to the meaning. Whole language proponents believe that language must be whole and meaningful because the interrelationships of the three language systems cannot be discovered or used by the learner in any other way.

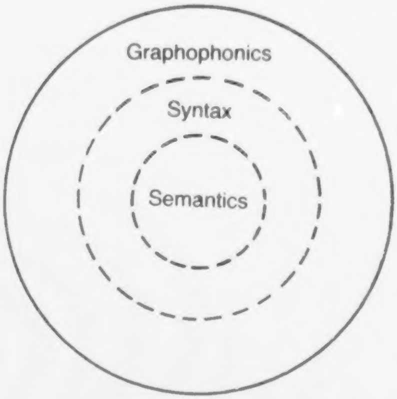


Figure 5. The whole language model.

Reader strategies reflecting a whole language model. Readers who hold whole language models of reading focus on meaning and large chunks of language during *The Reading Interview* (Burke, 1978).

> Interviewer: When you are reading and you come to something you don't know, what do you do?
>
> Carolyn: I do a story—read it to somebody. I make up something that makes sense. Sometimes I guess or sound it out.
>
> Interviewer: Do you ever do anything else?
>
> Carolyn: Get another book.

Carolyn's focus on meaning is clear. While she recognizes "sounding out" as an available strategy, she discusses it within the context of making a story make sense.

The remainder of Carolyn's interview is consistent with the first question. When asked how she would help someone having difficulty with reading, her preferred strategy was to "read the book for them—then they would read it." Her teacher was perceived as having a meaning focus also; students in her classroom are told to go on reading when they come to something they do not know. Carolyn intuitively realizes the importance of keeping language whole and meaningful when reading.

Carolyn, a first grader, was not part of the study described in these pages.[1] The 13 first graders in this study, of whom Shanna and Shane were two, focused on letter-sound relationships and words. Most of the children's theoretical orientations were located toward the phonics end of the model continuum; most of the children's interviews closely resembled Shane's interview.

Texts reflecting a whole language model. Meaning is the primary language focus of whole language basal reader programs. Because the authors believe that the language systems (graphophonics, syntax, semantics) are interrelated, they provide children with messages that convey meaning no matter which language system is being focused on instructionally. Language is not pulled apart—it is kept whole so that children can discover how cue systems work in their natural interrelationships.

Texts are constructed or selected with the reader's knowledge of language and of the world in mind, and with the belief that by reading, children will discover the regularities of print. Because children are believed to actively contribute their own knowledge in the process of reading, it is thought that the children will find easiest to read those stories that are written about something that is highly familiar to them in natural and familiar language. Difficult words and complex sentences are not avoided if they are a natural part of a good story; enough context is provided in the story to support children in their efforts to use effective strategies as they construct the author's meaning.

Those who hold a whole language model of reading instruction believe that the traditional notion of readability should be replaced by the notion of predictability. Readability formulas are based on textual factors; readers are only peripherally taken into consideration in determining vocabulary difficulty and grammatical complexity. Predictability, on the other hand, is a concept that takes into account the interface between reader and text. The better the match between the knowledge of the reader and the language and content of the text, the more predictable the reader should find the author's message and the form of the message. Goodman (1972) supports such a notion: "A given sequence will be easier to read to the extent that what the reader is most likely to predict actually occurs" (p. 153).

---

[1] The author thanks Diane DeFord (1978) for providing this interview.

112

Because the authors of Scott, Foresman's *Reading Systems* believe that children should be given the opportunity to "use the cues and develop the comprehension strategies (predicting, confirming, comprehending) most useful to them" (Manual, 1971, p. 7), they recommend that teachers should "always keep. . . help to a minimum" (p. 169) when giving the children a story to read. Words are not introduced prior to a story. If the story is truly predictable, the children should be able to make good predictions as to language and content, including words that they may have difficulty with in a less supportive context. Harste and Burke (1978) note that:

> Words are not either known or unknown. We simply must abandon notions of empty cells and full cells when it comes to words. Each word has features of meaning attached to it which have developed over time and through encounters in our world. (p. 32)

Repetitive and cumulative syntactic and semantic sequences are often characteristic of predictable stories. Such repetition provides even more redundancy than language typically supplies; if language cues are not effectively used the first time they appear in a story, the repetitive or cumulative context should ensure better predictions on the child's part as the cues reappear.

"The Three Little Pigs," a story from *Reading Systems*, was chosen for inclusion in this study primarily because of the familiarity of the story to children. A child's familiarity with a story ensures a close match between the child's knowledge of the world (in this case, specific knowledge of the story) and the author's presentation.

It was assumed that the language of the author and reader were also closely matched in this story; the sentence structures appeared to be natural and well-formed. The language is also rendered predictable because the story is structured with repetitions of episodes and repetition of language structures between episodes. Because the language units that are repeated are so large, the reader is referred to the entire text in Table 1. An excerpt from the text would not do it justice.

"The Great Big Enormous Turnip" (Tolstoy, 1971), also from *Reading Systems*, was chosen for inclusion in the study primarily because of its cumulative and repetitive sequences. Each story episode is a repetition of the previous episode with the addition of a character to the previous cast of characters. Like "The Three Little Pigs," the language structures are the same within each episode.

An excerpt from "The Great Big Enormous Turnip" provides an example of the cumulative pattern and sentence repetition:

> He called the old woman.
> The old woman pulled the old man.
> The old man pulled the turnip.
> They pulled—and pulled again.
> But they could not pull it up.
> So the old woman called her granddaughter.
> The granddaughter pulled the old woman.

Table 1
The Three Little Pigs[a]

1. Once upon a time there were three little pigs.
2. One day the three little pigs left home.
3. Each little pig wanted to build a house.
4. The first little pig made a house of straw.
5. The second little pig made a house of sticks.
6. The third little pig made a house of bricks.
7. A big bad wolf saw the little pigs.
8. "A little pig will make a good lunch for me," he said.
9. The big bad wolf went to the house of the first little pig.
10. The wolf called, "Little pig, little pig, let me come in."

11. "Not by the hair of my chinny chin chin!" said the first little pig.
12. "Then I'll huff and I'll puff and I'll blow your house in!" called the wolf.
13. So he huffed and he puffed and he blew the house in.
14. But the first little pig got away.
15. He ran into the woods to hide.
16. The big bad wolf went to the house of the second little pig.
17. The wolf called, "Little pig, little pig, let me come in."
18. "Not by the hair of my chinny chin chin!" said the second little pig.
19. "Then I'll huff and I'll puff and I'll blow your house in!" called the wolf.

20. So he huffed and he puffed and he blew the house in.
21. But the second little pig got away.
22. He ran into the woods to hide.
23. The big bad wolf went to the house of the third little pig.
24. The wolf called, "Little pig, little pig, let me come in."
25. "Not by the hair of my chinny chin chin!" said the third little pig.
26. "Then I'll huff and I'll puff and I'll blow your house in!" called the wolf.
27. So he huffed and he puffed, and he huffed and he puffed.
28. But the wolf couldn't blow the house in.

29. The wolf got on top of the house.
30. He called, "Little pig, little pig!
31. I'll come down the chimney and eat you up!"
32. The third little pig called, "Come on down."
33. The third little pig took the cover off the pot of water.
34. The water was very hot.
35. The wolf slid down the chimney.
36. He landed in the water with a big splash.
37. And that was the end of the big bad wolf.

---

[a] "The Three Little Pigs" from *Scott Foresman Reading Systems, Level 3*, 1971. Reproduced with permission.

The old woman pulled the old man.
The old man pulled the turnip.
And they pulled—and pulled again.
But they could not pull it up.
The granddaughter called. . . . (pp. 10-18)

In this story, the sequence of characters from biggest to smallest should be easily identifiable by the reader and should become a semantic cue as the reader catches on to the semantic structure the author has provided. Likewise, for any child who is familiar with "The Farmer in the Dell," the sequence of the specific characters should be highly predictable.

Beginning Readers' Responses to More and Less Predictable Texts

The following presentation of the current study is divided into four parts: (a) administration of the study, (b) data collection and analysis, (c) results, and (d) discussion.

Administration of the Study

Thirteen first graders, eight boys and five girls, were randomly selected from the middle reading groups of five first grade classrooms located in the same working-class neighborhood school of a small midwestern town. The children had all been attending first grade for seven months.

Three of the children produced fewer than 25 miscues on more than one of the four stories read for the study and were excluded from the data base for all miscue analysis. For that reason, miscue findings are based on 10 children; all other analyses include data from 13 children.

It was determined that the 13 children held theoretical models of reading that focused on letter-sound relationships and/or words. Most of the models were located toward the phonics end of the model continuum.

Each child read the four texts described in the previous section on reading models. "The Three Little Pigs" and "The Great Big Enormous Turnip" were selected to represent a whole language model of reading and, in this study, are referred to as "more predictable." That is, it was expected that the children would be more able to predict the authors' language and content and would be more able to integrate their knowledge about the stories with the authors' knowledge. The Spache readability formula (Spache, 1974) designated a 2.6 grade level for "The Three Little Pigs" and a 1.6 grade level for "The Great Big Enormous Turnip."

"The Big Stand" and "Pork Chops and Applesauce" were selected to represent phonics and skills models of reading, respectively. Both models share the assumption that meaning must be accessed through word recognition. Stories based on such an assumption consist of restricted vocabulary and stilted lan-

115

guage patterns and as such, are rendered "less predictable." For this reason, it was expected that the children would be less able to predict the authors' language and content in "The Big Stand" and "Pork Chops and Applesauce." According to the Spache readability formula, "The Big Stand" was written at a 1.3 grade level; "Pork Chops and Applesauce" was written at a 2.0 grade level.

Following *Reading Miscue Inventory* (Goodman, Y., & Burke, 1972) procedures, each child orally read four texts in two sessions held on succeeding days. Both sessions were tape-recorded in their entirety. Seven of the children read the more predictable texts the first day; six read the less predictable texts the first day. Prior to reading the stories, the children were given the following instructions: "If you come to something you don't know when you're reading, skip, guess, or do whatever you usually do when you're reading alone. I won't be able to help you." Following the oral reading of each story, the children retold it.

At the beginning of the first session, *The Reading Interview* (Burke, 1978) was administered to determine reader theoretical orientation. At the end of the second session, each child was asked to rank the four stories according to (1) how well the stories were liked and (2) how easy the stories were to read.

Data Collection and Analysis

Several types of data were collected: (1) miscue data, (2) retellings, (3) verbalizations made in the process of reading, and (4) story preferences. All data, with the exception of story preferences, were analyzed by two coders. Differences in judgment, when they existed, were resolved after discussion between coders.

Miscue data. Miscue data for 10 children were analyzed in several ways. Using *Reading Miscue Inventory* (RMI) (Goodman, Y., & Burke, 1972) procedures, each child's utilization of the graphophonic, syntactic, and semantic cuing systems was examined as well as the number of miscues per hundred words and number of corrections. All miscues produced by the readers were analyzed. Using the procedures of the *Reading Miscue Inventory: Evaluation Form* (Goodman, Y., & Burke, in press), a judgment was made of the semantic acceptability of every sentence produced by the readers. A summary of the miscue data collected and the analysis procedures utilized for each measure are specified in Table 2.

Retellings. Retellings were transcribed and analyzed according to RMI procedures. A retelling score for the narrative material in this study was made up of 30 points for character recall and development, 30 points for events recall, and 20 points each for plot and theme statements, for a total of 100 points. The children's retellings were compared with retelling outlines formulated by the author.

Verbalizations. Eleven of the 13 children produced spontaneous verbalizations in the process of orally reading. Those protocols were categorized according to their appearance in the process of reading a more or a less predictable story and each verbalization was categorized as to its focus: (1) meaning, (2) a word, or (3) the situation surrounding the reading or research task.

116

Table 2
Miscue Data Collection and Procedures

Type of Miscue Data	Analysis Procedures
Graphophonic cue utilization	[(high + some levels of graphic similarity) + (high + some levels of sound similarity)] ÷ 2 (sound category and graphic category)
Syntactic cue utilization	strength + partial strength + overcorrection levels of grammatical relationships
Semantic cue utilization	high + some levels of comprehension pattern
Semantic acceptability of sentences	A yes or no judgment was rendered for this question for every sentence: Is the sentence (as last produced by the reader) semantically acceptable within the context of the story? The number of yes judgments ÷ the number of sentences in the story or story quarter.
Corrections	The number of complete miscue corrections in the story or story quarter ÷ the total number of miscues in the story or story quarter.
Miscues Per Hundred Words (MPHW)	The number of miscues in the story or story quarter ÷ the number of words in the story or story quarter x 100.

Story preferences. Each child ranked the four stories from easiest to hardest and from best liked to least liked. The rankings were combined to provide a preference scale that ranged from 2 (most preferred story) to 8 (least preferred story).

Results

Because all the subjects in this study were judged to be average first grade readers and because they consistently held models of reading which fell between the phonics and skills points on the model continuum, only group data (means) on the children's interactions with more and less predictable texts will be presented. The focus, in other words, was on the effect that more and less predictable texts had on the children's reading behaviors.

Before turning to miscue findings, it should be noted that the miscue analyses were based on a total of 2,462 miscues made by 10 first grade readers in the process of reading four texts. The number of miscues made in each text is listed in Table 3.

Table 3
Total Miscues Per Text

Stories	Number of Miscues
"The Three Little Pigs"	475
"The Great Big Enormous Turnip"	481
"The Big Stand"	399
"Pork Chops and Applesauce"	1,107
Total	2,462

Language cue system utilization. All readers use three language cue systems as they read: the graphophonic, syntactic, and semantic systems. Proficient readers use the three systems in a balanced, interrelated fashion at approximately a 75% level. That is, proficient readers can be expected to show a moderate but even utilization of the three language systems (Goodman, Y., & Burke, 1972).

Figure 6 displays the utilization of the three cuing systems for miscues produced by the group of first graders across all four stories. As was expected from the children's theoretical orientations to reading (a focus on words and/or relationships), the first graders more heavily used the graphophonic cuing system across all four stories than they did the semantic or syntactic cuing systems.

Graphophonic cue utilization ranged from 75-82% on the four stories, while syntactic cue utilization ranged from 31-56%, and semantic utilization ranged from 25-53%. In the process of reading all four stories, the children focused more attention on producing responses that graphophonically resembled expected responses than they did on producing responses that made sense and sounded like language. An excerpt from a reading of "Pork Chops and Applesauce" provides an example of one child's attention to the graphophonic cuing system:

Text	Student's version of text
Then Ted remembered something	Then *Toad* .... something.
He liked applesauce with pork chops.	He liked *airplanes airplanes* with pork *chips*.
Sometimes Mother forgot to get the applesauce.	*Someth–Sometimes* Mother .... to *go go get* the *airplane*.

The miscues in this excerpt were typical of this group of children, in that the child produced omissions and highly graphophonically similar substitutions. Both the child's substitutions and omissions were often syntactically unacceptable and even more often, semantically unacceptable.

The children's theoretical orientations had a limiting and pervasive influence on reading strategies as predicted earlier; all 10 children focused on the graphophonic system and did not effectively use the syntactic and semantic systems while reading. The two main strategies verbalized in *The Reading Interview* by the children for dealing with something unknown were sounding the word out and asking someone what the word was. The high graphophonic similarity of the word substitutions is evidence of the children's "sounding out" attempts. Omissions are evidence of the children's notions about "known" and "unknown" words. They decide whether they know or do not know each word; if it is "known" the expected response or a graphophonically similar substitution is produced and if it is "unknown," the word is often omitted. If the researcher had not given explicit instructions that help would not be provided, the children would probably have asked for the words rather than omitting them in many cases.

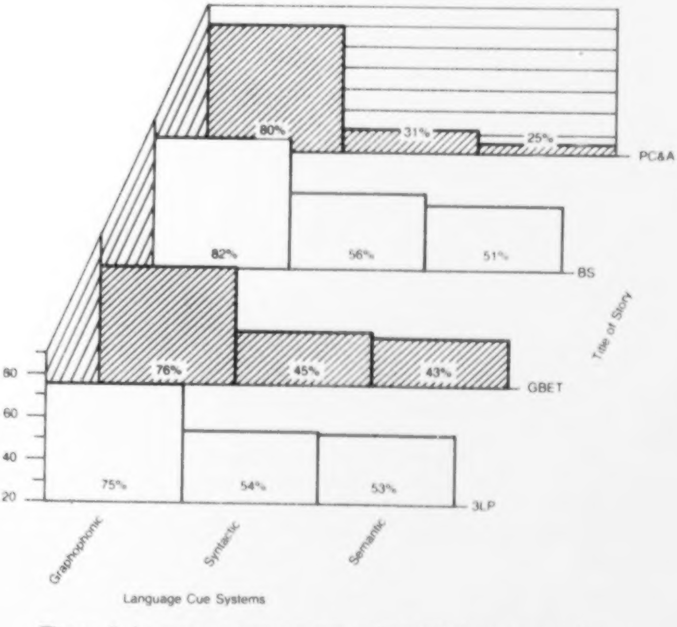


Figure 6. Language cue system utilization/story comparison.

Figure 6 also reveals that different stories had different effects on the children's reading strategies; variations in print setting (in this case, variations in texts) did affect the reading process. "The Three Little Pigs" and "The Big Stand" were most supportive in encouraging the children to use their knowledge of the syntactic and semantic systems while reading. "The Great Big Enormous Turnip" was less supportive, and "Pork Chops and Applesauce" was quite unsupportive in encouraging use of the syntactic and semantic systems.

Some of these findings were puzzling. "The Big Stand" was selected as one of the less predictable stories, yet it was found to be one of the most supportive in encouraging syntactic and semantic system usage. "The Great Big Enormous Turnip," selected as one of the more predictable stories, was less supportive than expected. The data were puzzling for another reason: The children appeared to become more effective readers (according to this researcher's judgment, formulated as the children read) in some sections of the predictable stories. These intuitions led to an examintion of cue system utilization within each quarter of each story. If the readers were, in fact, more effective users of the syntactic and semantic cue systems in some parts of the stories, an analysis could be expected to reveal variations in cue utilization within stories.

For such an analysis, each story was divided into four parts at the sentence level. For example, because "The Great Big Enormous Turnip" had 44 sentences, each quarter consisted of 11 sentences. Cue system utilization analysis of this story was done on the 11 sentences in each quarter. Because only substitution miscues are marked as to graphophonic similarity and because a large number of omissions considerably lowered the number of substitution miscues which could be considered in each quarter, a decision was made to examine only syntactic and semantic cue system utilization by quarters. The results of syntactic and semantic system analyses across quarter stories appears in Figure 7.

Clear variations in levels of cue utilization between story quarters are apparent in Figure 7. For example, syntactic cue system utilization in the first quarter of the stories was highest in "Pork Chops and Applesauce" (55%); for that story it then dropped to 30%, 33%, and 24% for the second, third, and fourth quarters, respectively. Semantic cue utilization in "Pork Chops and Applesauce" was only 39% for the first quarter of the story but dropped in the same way as syntactic cue utilization: to 25%, 28%, and 17% for the final three quarters of the story.

The variations in cue utilization levels of "The Great Big Enormous Turnip" were in stark contrast to those in "Pork Chops and Applesauce." Instead of declining as they do in "Pork Chops and Applesauce," semantic and syntactic cue utilization levels continually rise in "The Great Big Enormous Turnip" (except for a 7% drop in syntactic use in the last quarter). In fact, semantic and syntactic cue utilization for the first quarter of "The Great Big Enormous Turnip" is lowest of the stories (27% syntactic; 26% semantic) and in the fourth quarter, the highest of the stories (55% syntactic; 65% semantic).

In both syntactic and semantic cue utilization, "The Three Little Pigs" increased through the third quarter when reading effectiveness was the highest for all stories in any quarter (78% syntactic; 76% semantic). Cue utilization then plummets in the fourth quarter to 48% syntactic and 44% semantic cue utilization, slightly lower than observed in the first quarter.

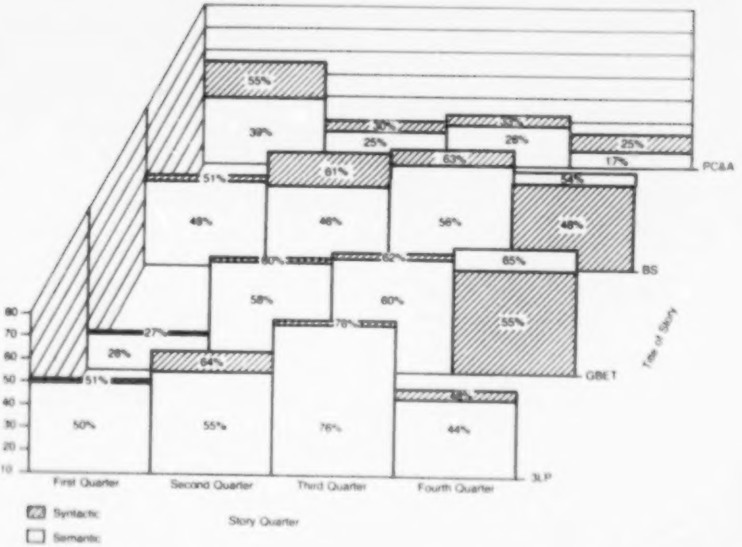


Figure 7. Syntactic and semantic cue system utilization/story quarters.

"The Big Stand" shows the most even performance of all the stories in syntactic and semantic cue utilization. Syntactic cue utilization ranged from 48% to 63%; semantic cue utilization ranged from 46% to56%. The difference between the lowest and highest cue utilization across any two quarters was only 15%. Syntactic cue utilization was a percentage point higher than that of "The Great Big Enormous Turnip" for the second and third quarters but semantic cue utilization was lower for both quarters.

The two more predictable stories, in this quarter-story analysis, clearly are supportive of the readers' strategies in using semantic and syntactic cues in some story quarters. The readers are encouraged by the predictability of the texts to use increasingly more effective strategies in the second and third quarters of "The Three Little Pigs" and in the second, third, and fourth quarters of "The Great Big Enormous Turnip."

The two less predictable stories had very different effects on cue utilization of the readers. Cue utilization in "The Big Stand" resembled that of the more predictable stories; it increased in some quarters and syntactic cue utilization was at a level similar to that of "The Great Big Enormous Turnip." Though the story encouraged a more effective utilization of syntactic cues, the children's utilization of semantic cues did not increase appreciably. In "Pork Chops and Applesauce," on the other hand, cue utilization not only decreased but was at a very low level; the story discouraged effective reading strategies.

Semantic acceptability of sentences. Degree of semantic acceptability of miscues and correction behavior are two important indicators of reading com-

Table 4
Semantically Acceptable Sentences: Story Comparison

Stories	Percent Semantically Acceptable
"The Three Little Pigs"	67%
"The Great Big Enormous Turnip"	61%
"The Big Stand"	62%
"Pork Chops and Applesauce"	28%

prehension (Carey, 1978; Goodman, K., & Burke, 1973). An effective reader is one who monitors the sense of the message that is being constructed.

Semantic acceptability of miscues and corrections are the two factors that constitute the analysis of semantic acceptability of sentences. In order to determine the level of semantic acceptability of sentences, each sentence was read as last produced by the reader (including corrections), and a judgment was made as to whether the sentence was semantically acceptable within the context of the total story (Goodman, Y., & Burke, in press).

The percentage of semantically acceptable sentences for each story is presented in Table 4.

There appear to be no real differences between the first three stories in the level of semantically acceptable sentences produced by the first grade readers. Between 61% and 67% of the sentences for those three stories were semantically acceptable. On the other hand, the readers produced only 28% semantically acceptable sentences while reading "Pork Chops and Applesauce."

Figure 6 is an analysis of the percentage of semantically acceptable sentences within each quarter story. Once again, quarter story analyses revealed variations in reading effectiveness that story level analyses did not. A comparison of "The Three Little Pigs" and "The Big Stand" shows that the children produced 53% semantically acceptable sentences in the first quarter of both stories and 63% and 65%, respectively, in the last quarter of each story. The children produced 13% more semantically acceptable sentences in the second quarter of "The Big Stand" than they did in the first quarter but then leveled off for the other quarters.

In "The Three Little Pigs," on the other hand, the children's sentences changed from 53% semantically acceptable sentences in the first quarter to 73% for the second quarter and 84% for the third quarter. Until the last quarter, the children became more and more effective readers as they progressed through the story. In the last quarter of "The Three Little Pigs," children's sentences declined to a 63% semantic acceptability rate, lower than the second and third quarters but higher than the first quarter.

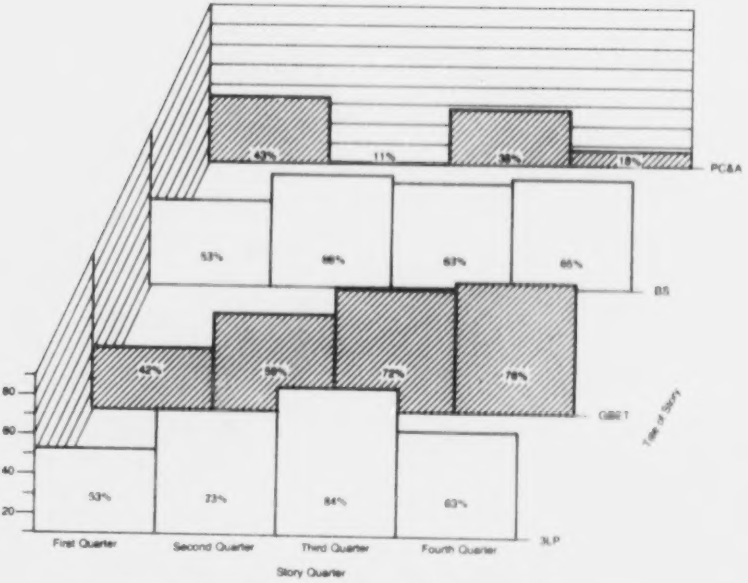


Figure 8. Semantically acceptable sentences/story quarters.

Levels of semantic acceptability in "The Great Big Enormous Turnip" made a steady progression upward (42%—59%—72%—76%). The last quarter of the story was read more effectively than the last quarter of any of the other three stories.

The level of semantically acceptable sentences in the first quarter of "Pork Chops and Applesauce" was 43%, about the same as "The Great Big Enormous Turnip." However, the level of semantically acceptable sentences in the second quarter rose in "The Great Big Enormous Turnip," while it dropped dramatically to only 11% in "Pork Chops and Applesauce." Although the semantic acceptability of sentences rose again in the third quarter of "Pork Chops and Applesauce" to 38%, it once again dropped sharply in the fourth quarter to 18%.

The same conclusion can be reached for the semantic acceptability of sentences as for syntactic and semantic cue utilization at the quarter story level; some quarters of the two more predictable stories clearly encouraged more effective reading. As the children progressed through "The Great Big Enormous Turnip" they read more effectively, and though there was a drop in effectiveness in the fourth quarter of "The Three Little Pigs," the children became increasingly more effective until that point in the story. In fact, the third quarter of "The Three Little Pigs" encouraged the most effective reading of any story quarter.

The less predictable stories affected reading effectiveness in different ways. The second quarter of "The Big Stand," like the more predictable stories, showed that the children become more effective readers but the remainder of the story did not encourage a further increase in reading effectiveness. "Pork Chops and Applesauce," on the other hand, had the effect of discouraging meaning construction. Neither story supported effective reading strategies to the same extent as the more predictable stories.

Miscues per hundred words and corrections. The level of semantic acceptability of the sentences produced by readers is dependent on the interaction between the semantic acceptability of miscues in the sentence and correction behavior. It also can be affected by the number of miscues per hundred words. The section that follows examines the findings related to correction behavior and miscues per hundred words (MPHW).

Figure 9 is a comparison of miscues per hundred words and semantic acceptability of sentences. In general, the two measures varied in opposite directions: as MPHW decreased, semantic acceptability of sentences increased. The relationship is consistent except in (a) "The Big Stand" where the semantic acceptability of sentences and MPHW both rose for the second quarter and then remained level for the remainder of the story, and (b) in the fourth quarter of "The Great Big Enormous Turnip" where both MPHW and semantic acceptability of sentences rose slightly. MPHW range from a low of 6.4 in the third quarter of "The Three Little Pigs" (where semantic acceptability of sentences was the highest at 84%) to a high of 29.7 in the second quarter of "Pork Chops and Applesauce" (where semantic acceptability of sentences was the lowest at 11%). Story level miscue acceptability (see Table 5) for the two more predictable stories was lower than MPHW for the two less predictable stories.

Figure 10 is a comparison of the percent of miscues that were corrected and the semantic acceptability of sentences. In general, the two measures co-varied: as correction behavior increased, the semantic acceptability of sentences increased. This relationship was consistent except for the fourth quarters of both "The Big Stand" and "The Great Big Enormous Turnip" where correction behavior decreased while levels of semantic acceptability increased.

Table 5
Miscues Per Hundred Words: Story Comparison

Stories	Miscues Per Hundred Words
"The Three Little Pigs"	10.3
"The Great Big Enormous Turnip"	14.3
"The Big Stand"	18.6
"Pork Chops and Applesauce"	25.8

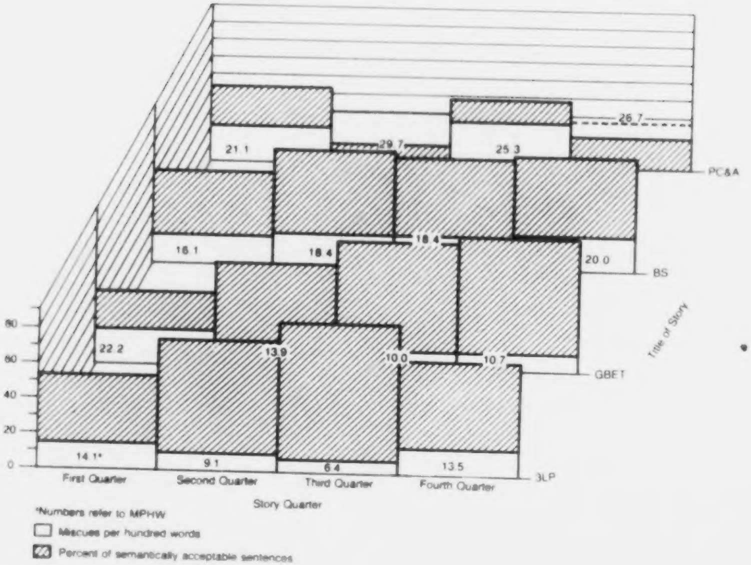


Figure 9. A comparison of MPHW and semantic acceptability of sentences/story quarters.

Clearly, a high level of semantic acceptability of sentences, though not always dependent on MPHW and correction levels, was related to a lower level of MPHW and a higher correction level for these first graders (Goodman, Y., 1967). For example, in the most predictable quarter of all stories, the third quarter of "The Three Little Pigs," the children made fewer MPHW (6.4) and more corrections (46%) than in any other quarter, and used the semantic and syntactic cue systems more effectively (76% and 78%, respectively) than in any other quarter. All of this contributed to the highest level of semantic acceptability of sentences within any quarter—84%. By the third quarter of "The Three Little Pigs," the children were reading quite effectively and efficiently. In 84% of the sentences, they either predicted exactly what the author had to say or their miscues kept the author's intended meaning intact when their predictions were not exact.

An example of a quarter in which the level of semantic acceptability did not vary in the same direction as corrections and in the opposite direction of MPHW was the fourth quarter of "The Great Big Enormous Turnip." In this quarter, even though the level of corrections decreased from the previous quarter (from 27% to 18%) and MPHW increased slightly (from 10.0 to 10.7), semantic acceptability of sentences also increased (from 72% to76%). This discrepancy was due to the children's increased effectiveness in using the semantic cuing system in the fourth quarter (a rise of 5% between the third and fourth quarters). In other words, although the children made slightly more miscues and corrected fewer of their miscues, their miscues were more semantically acceptable. In the fourth quarter of "The Great Big Enormous Turnip," the children's predictions did not exactly

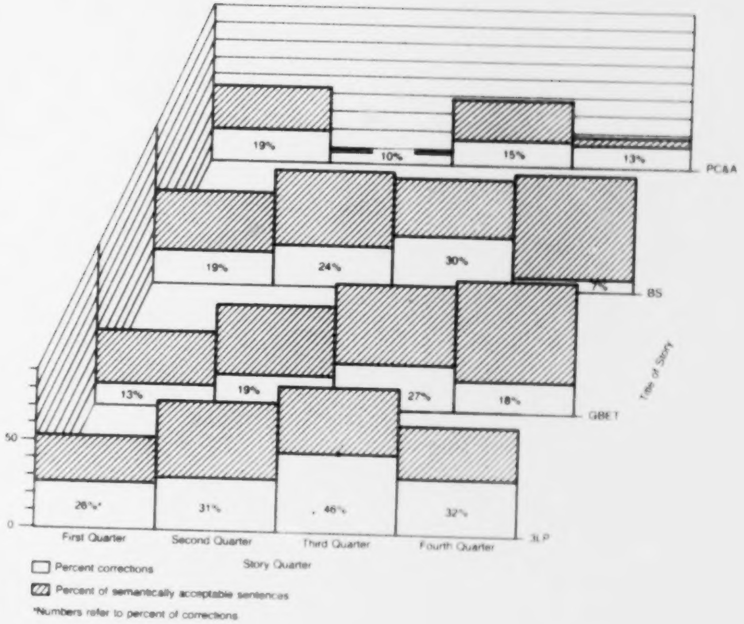


Figure 10. A camparison of percent of corrections and semantic acceptability of sentences/story quarters.

match the text as often as they did in the third quarter of "The Three Little Pigs," but their predictions kept the author's intended meaning intact in 76% of the sentences (in the fourth quarter of "The Great Big Enormous Turnip"). The reading profile observed in this quarter is much like that observed in older children (Goodman, K., & Burke, 1973).

Retellings. A retelling score is a reflection of the reader's integration of his own knowledge and the author's intended meaning. It is said that the scores are only a "reflection" of the reader's integration of his own knowledge and the author's intended meaning because it is unlikely that the reader will retell everything he knows about the story. That may be particularly true when the reader retells the story to a person who has just listened to the story being read, as was the case in this study. Such a situation is not natural—the child is likely to leave out information because he is confident that the listener knows the story. The mean retellings (of 100 possible points) for the four texts are given in Table 6.

The children were able to generate more information that matched the author's intended meaning after reading the more predictable stories than they were after reading the less predictable stories. Some of this information, particularly in retellings of a well-known story like "The Three Little Pigs," was generated from previous knowledge the children held of concepts in the story, or about the story

Table 6
Retelling Scores: Story Comparison

Stories	Retelling Scores
"The Three Little Pigs"	78.7
"The Great Big Enormous Turnip"	60.8
"The Big Stand"	51.8
"Pork Chops and Applesauce"	23.9

itself. The score does not represent only a simplistic notion of meaning gain as if the reader were a blank slate upon which the author's meaning must be written.

For example, it is not valid to claim that the readers gained 78.7% meaning from "The Three Little Pigs." They gained some proportion of that 78.7% from reading the story; the remainder consists of the reader's knowledge that was brought to the story and integrated with the author's intended meaning. The same evaluation can be made of the score from each story, though perhaps not to the same degree because of the high level of familiarity that the children had with "The Three Little Pigs." Nevertheless, all retellings are partially a reflection of the knowledge the reader brought to the reading and retelling situation.

A specific example of children bringing their own knowledge to the reading situation occurred in "The Great Big Enormous Turnip." In reading the story the children produced a variety of substitutions for "turnip" and "granddaughter" or omitted them. However, in the retelling of the story, most children referred to the turnip as a plant and to the granddaughter as a girl. Both of these labels were generated from the children's previous knowledge of the world and though they did not identify the author's exact label or meaning, they preserved the essential meaning of the story in their retellings through contribution of their own knowledge.

The notion of the reader's previous knowledge influencing retelling scores does not invalidate retelling scores as a measure of comprehension. It is a recognition, instead, of the importance of reader knowledge as a basis of comprehending strategies. The reader's ability to contribute knowledge to a story affects the reader's capability in predicting and confirming the author's language and content.

If the scores are considered to be reflections of the integration of the reader's and author's meanings, it can be concluded that the more predictable stories support both identification and integration of the author's meaning as well as the ability to predict at a more effective level. Both are important to reading as a process and as a product.

Verbalizations. Eleven of the 13 children produced spontaneous verbalizations in the process of reading the texts. That is, they made oral comments or asides that were not visually related to the text they were in the process of reading. Table 7 provides a listing in the left column of those verbalizations made in the process of reading the more predictable stories and a listing in the right column of the verbalizations made while reading the less predictable stories. The set of parentheses prior to each verbalization identifies the text being read at the time of the verbalization. When the verbalization is meaningless without additional information concerning the context in which it occurred, that information is also supplied.

Each verbalization was classfied as to meaning (M), word (W), or situation (S) focus, and is indicated in parentheses following each verbalization. If the child commented on a story event, predicted what might happen, summarized what had happened, made an inference, recognized the story organization, or brought his own knowledge or feelings to bear on the story, the verbalization was classified as having a focus on meaning. If the child commented on the difficulty of a word or made a comment related to a word recognition strategy, the verbalization was classified as having a word focus. Situation verbalizations included comments on when or where the story would end or the length of the story.

The meaning, word, and situation verbalizations were tallied; the results are presented in Table 8 as well as percentages that indicate the salience of each verbalization category.

The analysis revealed large differences in the focus of verbalizations spontaneously made as the children read the more and less predictable stories. Sixty-four percent of the verbalizations produced in the process of reading more predictable stories were related to the meaning of the stories, while only 7% related to meaning in the less predictable stories. Twenty-seven percent of the verbalizations in the more predictable stories were related to words, while 67% were related to words in the less predictable stories. Situation verbalizations were made 9% and 26% of the time in the more and less predictable stories, respectively. Many of the situation verbalizations in the less predictable stories had negative connotations; they consisted for the most part of concerns about where or when the story would end.

Other observations also revealed the kind and degree of the children's involvement with the stories. For example, it was noted several times that the children laughed and smiled in the process of reading the more predictable books; only one child had the same reaction (and then only once) in the process of reading a less predictable book. Several notes, instead, were made of children producing frustrated noises, looking to see where the story ended, and finding it difficult to sit still as they read the less predictable stories. On the other hand, one little boy became so involved in reading "The Three Little Pigs" that he read the pigs' and wolf's dialogues in different voices.

The focus of the children's verbalizations clearly reflected the theoretical orientations of the texts that they were reading. The less predictable books with letter-

Table 7
Verbalizations. More and Less Predictable Stories

More Predictable Stories

(GBET) Look what he's trying to do--pull out the leaf. (M)
(GBET) What they need is a gadget. (M)
(GBET) The end. Second story read. (S)
(3LP) He's going to trick him! (M)
(3LP) Cause he was cooked in it. (M)
(GBET) Still can't figure it (*turnip*) out. (W)
(GBET) I think that's *tornbita.* (points to *turnip*) (W)
(GBET) Turnip! That's what it is. (W)
(GBET) That's a long word!(W)
(3LP) Oh, that's *second*! (W)
(3LP) Oh, she's fat! (M)
(3LP) Oh, what a small window! (M)
(3LP) I wish I could make a house. (M)
(3LP) Oh, I've never seen a wolf do that. (M)
(3LP) I almost said *first little wolf*! [instead of *first little pig*]. (M)
(3LP) He's dancing. (M)
(3LP) This is a funny story. (M)
(3LP) If I were him, I wouldn't do that! (M)
(GBET) Don't get too big (she's talking to turnip). (M)
(GBET) Oh, this word again (*enormous*). (W)
(GBET) No wonder, he's so old. (M)
(GBET) I hate this (the word *granddaughter*). (W)
(GBET) What a short story. You can read it all in one day. (S) (Level 3)
(GBET) Where is the mouse? It's so small. (M)
(GBET) He's getting tired. (M)
(GBET) They should have pulled it out when it was like that (shows me little round circle with his fingers), right? (M)
(3LP) I can't read these big words. I can't read that word (*once*)

Less Predictable Stories

(PC&A) Now, I'm ready for the very last page. (S)
(PC&A) Is this where to stop? (S)
(PC&A) I never had that word (*grandmother*) before. (W)
(PC&A) If this didn't stand for Mr., it'd be m + r + r + r. (W)
(PC&A) Every time we get meat, there's blood all over it. (M)
(PC&A) I don't know that one either (*applesauce*). (W)
(PC&A) I don't know that word (*gingerbread*). (W)
(PC&A) Oh, I don't know that word (*gingerbread*) (W)
(PC&A) He's taking it all. (M)
(PC&A) This is the last page. There's a lot of words I'm skipping because I have to (W)
(BS) I had that word--I forgot it (*may*). (W)
(BS) Now, I had that word. If I say it...I forgot it...I'll get it wrong. I always count to 10...I counted to 10, but.... (W)
(BS) I keep getting mixed up on words because I read too fast. (W)
(BS) I wish I could read the whole thing (referring to whole book that BS is in). (S)
(PC&A) Is this where I stop? (S)
(PC&A) Do I read that? (referring to last page) (S)
(BS) Stop here? (S)
(PC&A) This is difficult. (W)
(PC&A) I can't read this. (W)
(PC&A) I don't know (for *remember*). (W)
(BS) There's that word again. (W)
(BS) This is difficult. (W)
(BS) There's that word again. (W)
(BS) Now, I'm going to try to sound it out instead of.... (W)
(BS) They got everything here (at the stand). (M)
(PC&A) Do I read this? (S)
(PC&A) Skip. I'll skip that over. I can't read it. (W)
(PC&A) Skip. (W)

Table 7 con't

and that word (*upon*). (W) It doesn't make sense. (M)
(3LP) I can't read all these words. (W)
(3LP) The brick one is safer. (M)
(3LP) That's not safe, is it? (M)
(GBET) I forget what them are called (points to turnip in picture). (M)
(3LP) Some of the story I'm memorizing. (M)
(3LP) Wwoo! (M)
(GBET) And boy did it grow! (M)
(GBET) Look how big that is! (M)
(GBET) Oh, now I got the picture! (M)
(GBET) Now, I know (*granddaughter*) (W)
(GBET) No--skip (*turnip*). (W)
(GBET) Oh, no! (said when cat called mouse). (M)
(3LP) Skip it--I don't know (*upon*) (W)
(3LP) Skip. (*build*) (W)
(3LP) I already like this one! (the story) (M)
(3LP) Skip. (*second* but then corrects) (W)
(3LP) He's smart. (the third pig) (M)
(3LP) This book is a cinch. (M)
(3LP) That's a new word I learned (*straw*). (W)
(GBET) That's the end. (S)
(GBET) Now, I lost my place. (S)
(GBET) I got a lot of pages to go! (S)
(GBET) There you go again (said when another character added to story). (M)
(GBET) Oh no, there it goes again (another character added). (M)
(GBET) They couldn't get it (the turnip)! (M)
(GBET) Oh no! (All characters are pulling each other.) (M)

(BS) Oh, no (said because he couldn't read *stand*). (W)
(BS) I skipped that word once. (W)
(BS) No. (She knows she didn't read word on page.) (W)
(BS) No. (She knows word is not *ER*.) (W)
(BS) Is this my last page, ain't it? (S)
(PC&A) No. (She knows word is not *ER*.) (W)
(PC&A) No. (She knows word is not *ER*.) (W)
(PC&A) No, I didn't read that. (W)
(PC&A) I don't know where to go up to (how far to read). (S)
(PC&A) Oh boy, it will be hard to finish this. (W)
(PC&A) One more minute (to work on a word). (W)
(PC&A) I don't know that word (*maybe*). (W)
(PC&A) That's the end. (Said after first two pages as a joke.) (S)
No further entries for verbalizations while reading the less predictable texts.

130

Table 8
Focus of Verbalizations: More and Less Predictable Texts

Focus	More Predictable Texts		Less Predictable Texts	
	Number	Percentage	Number	Percentage
Meaning	35	64%	3	7%
Word	15	27%	28	67%
Situation	5	9%	11	26%

sound and word focuses, encouraged the children to focus on words and finishing the stories. The more predictable books, with a meaning focus, encouraged the children to focus on meaning and enjoyment of the stories.

Story preferences. Each child ranked the four texts from best to least liked and from easiest to hardest to read. Best liked and easiest to read stories were coded with a one; least liked and hardest to read stories were coded with a four. The two scores were combined for a "story preference" score (see Table 9) that had a possible range from two for the most preferred story to eight for the least preferred story.

The children preferred "The Three Little Pigs" above the other three stories and ranked "Pork Chops and Applesauce" as least preferred. "The Great Big Enormous Turnip" and "The Big Stand" were equally preferred after "The Three Little Pigs." In explaining their rankings, many of the children commented on the "hard words" in "The Great Big Enormous Turnip." The words they identified as being difficult, "turnip" and "granddaughter," were repeated frequently in the story.

Table 9
Story Preferences

Story	Preference Rating	Spache Readability Level
"The Three Little Pigs"	3.2	2.6
"The Great Big Enormous Turnip"	5.2	1.6
"The Big Stand"	5.2	1.3
"Pork Chops and Applesauce"	6.9	2.0

Because of their theoretical orientations, the children worked repeatedly to produce responses which were graphophonically similar even though they had the knowledge (as noted in the retelling discussion) to make meaningful substitutions such as "plant" and "girl." A certain amount of frustration was certainly engendered in the process of repeatedly attempting these words and it undoubtedly affected the degree of story preference.

The Spache readability formula (See Table 9) did not predict well either the children's preferences or their actual reading performances. For example, the formula predicted that "The Three Little Pigs" would be the most difficult story. However, the children's reading effectiveness was at its height for "The Three Little Pigs" and all but one child chose it as the easiest to read.

The children apparently did not find that low frequency words and longer sentences made the stories in this study difficult to read. Their preferences, in fact, more closely agreed with the notion that more predictable books (which sometimes contain low frequency words and longer sentences) are easier to read. Further, the children's judgments of difficulty corresponded highly with miscue analysis and retelling indicators of difficulty.

Discussion

Both author's and readers' theoretical orientations to reading had pronounced effects on the encounters of the readers with the texts in this study. The first graders, who held models built on the premise of word recognition as a prerequisite to meaning construction, were observed to use the graphophonic system more than the syntactic and semantic systems in both the more and less predictable stories. Their stated focus on letter-sound relationships and words in *The Reading Interview* was, in fact, operationalized during the reading process.

However, the semantic structures of the texts, based on the author's theoretical models, modified the children's strategies. In spite of the pervasive and limiting effect of the children's perceptions of reading on their reading strategies, both process and product measures substantiated the expectation that the children would more effectively read the more predictable stories.

Retelling scores, reflecting both the information brought to the stories by the children as well as information gained from the stories, indicated higher levels of comprehension for the more predictable stories. An analysis of the children's spontaneous verbalizations indicated that the children focused on word recognition in the process of reading the less predictable stories while they focused on meaning in the process of reading the more predictable stories.

Whole story miscue analysis data did not entirely substantiate the expectation that the first graders would use more effective strategies when reading the more predictable stories. Syntactic and semantic cue utilization was similar for "The Three Little Pigs" and "The Big Stand"; it was lower for "The Great Big Enormous Turnip" and lowest for "Pork Chops and Applesauce." However, a quarter story miscue analysis of each story revealed that the more predictable stories did encourage more syntactic and semantic cue utilization in the process of reading some sections of the more predictable stories. A quarter story analysis of syntactic and semantic cue utilization as well as an analysis of semantic acceptability of sentences illustrated that a reader's strategies can be continuously modified by text structure in the process of reading. Degree of effectiveness of strategy utilization varied as much or more within some stories as across stories. Retrospective examinations of the text structures produced viable reasons for such variations in performance, all based on the notion of predictability.

As a general rule which has been previously observed (Menosky, 1971), the children read more effectively after the beginning of the story, in this case, after the first quarter of the story. It can be surmised that the children were able to effectively modify their strategies after sampling the author's language and content in the first section of the story. That general observation did not hold for "Pork Chops and Applesauce," a less predictable text structure that discouraged the children's utilization of syntactic and semantic cues after the first quarter. Due to the inadequacy of the semantic text structure in "Pork Chops and Applesauce," the children were unable to effectively predict the author's content or language in the final three quarters of the story.

The more predictable stories were the only texts in which effective cue utilization was observed to significantly increase after the second quarter. The children used increasingly more effective strategies through the final quarter of "The Great Big Enormous Turnip." The cumulative and repetitive pattern in the structure of the text permitted better reader predictions of the author's language and content. Likewise, effective cue utilization increased in "The Three Little Pigs" through the third quarter of the story; the syntactic and semantic repetitions in the second and third quarters encouraged better predictions.

As alluded to earlier, however, the "more predictable" stories were proven to be "less predictable" in some quarters. Extremely poor strategy utilization was observed in the first quarter of "The Great Big Enormous Turnip"; a retrospective examination of that part of the text structure revealed that words were repeated without any underlying semantic basis, resulting in stilted language and little content. The message in the entire first quarter of the story could have been succinctly stated in a sentence or two.

For very different reasons, the fourth quarter of the "Three Little Pigs" was less predictable as well. In this case, the pattern established in the story is broken in the fourth quarter; the syntactic and semantic repetitions of the second and third quarters cease as the final episode begins in the fourth quarter. The final episode contains information that is different from that appearing in several previous episodes; the children revert to using less effective strategies very similar to those utilized in the first quarter of "The Three Little Pigs," the beginning of the story. They find themselves not as able to predict the author's language and content in the ending of the story.

A quarter story analysis of "The Big Stand," selected as a less predictable story, showed a comparatively even performance after an increase between the first and second story quarters. The children did not use more effective strategies as they did in both of the more predictable stories; nor did they use less effective ones as they did in "Pork Chops and Applesauce." A retrospective examination of the text structure revealed that, although the content and language were stilted, the author had used repetitive semantic sequences similar to those found in more predictable stories. Those sequences, though not stated in the same language each time, undoubtedly contributed to the predictability of this story when compared with that observed in "Pork Chops and Applesauce."

It can be surmised, then, that the variations noted in reader responses to texts within the stories themselves can be traced to factors of predictability, as can the variations across stories. In fact, such within story variations only lend more credence to the notion of predictability. The reader/text encounter is a unique one not only from story to story but within a story as well. Reader performance can sometimes be observed to be positively cumulative as in "The Great Big Enormous Turnip" and in the first three quarters of "The Three Little Pigs," while sometimes reader performance decreases as in the first quarter of "The Great Big Enormous Turnip" and the last three quarters of "Pork Chops and Applesauce," and only slightly variable as in the last three quarters of "The Big Stand." And, as in the case of the fourth quarter of "The Three Little Pigs," readers can be observed to use strategies similar to those used as they began a story.

It should be noted here that, given different reader perceptions of reading, the observed responses of the readers to the texts might have been quite different. For example, a group of first graders who employed a whole language model with a focus on meaning might have produced different process and product data in response to the four texts. Both text structure and reader perception can be expected to produce variations in reader response to texts (see Figure 1).

Clearly, the notion of predictability should be considered as residing on a continuum dependent on both the reader and the text. For first graders, factors like repetition of large chunks of language and episodes, familiarity of story lines, story sequences, or concepts appear to encourage more effective reading. However, these factors and others vary uniquely and to different degrees from text to text and from reader to reader. Though some of the factors have been found to be measurable (Rhodes, 1979), the unique combination of the number of factors that influence readers' responses to texts, render attempts to describe the transaction quite complex.

In spite of the complexity involved, it is clear that attempts to choose texts for students on the basis of potential predictability is a sound process and one firmly based on current reading theory postulating transaction between reader and text. Choosing texts for students based on a traditional notion of readability does not adequately take into account the role of the reader in the reading process and, in fact, is not based on theory and data which show the reader to be an active predictor and confirmer of text.

References

Burke, C. L. *The reading interview*. Unpublished guide, the Reading Program, Indiana University, 1978.

Carey, R. F. *A psycholinguistic analysis of the effects of the semantic acceptability of oral reading miscues on reading comprehension.* Unpublished doctoral dissertation, University of Connecticut, 1978.

Clymer, T., & Jones, V. W. Pork chops and applesauce. In *Seven is magic.* From Reading 360. Lexington, Mass.: Ginn, 1973. (a)

Clymer, T., & Jones, V. W. *Teacher's edition: Seven is magic.* From Reading 360. Lexington, Mass.: Ginn, 1973. (b)

DeFord, D. E. *A validation of an instrument to determine a teacher's theoretical orientation to reading instruction.* Unpublished doctoral dissertation. Indiana University, 1978.

Goodman, K. S. The reading process: Theory and practice. In R. E. Hodges & E. H. Rudorf (Eds.), *Language and learning to read.* New York: Houghton Mifflin, 1972.

Goodman, K. S., & Burke, C. L. *Theoretically based studies of patterns of miscues in oral reading performance.* Final report, Project No. 9-0375, Grant No. OEG-0-9-B20375-4269. U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research, April 1973.

Goodman, Y. M. *A psycholinguistic description of observed oral reading phenomena in selecting young beginning readers.* Unpublished doctoral dissertation, Wayne State University, 1967.

Goodman, Y. M., & Burke, C. L. *The reading miscue inventory: Evaluation form.* In press.

Goodman, Y. M., & Burke, C. L. *The reading miscue inventory: Procedure for diagnosis and evaluation.* New York, Macmillan, 1972.

Harris, T. L., Creekmore, M., & Greenman, M. H. The big stand. In *The sun tree.* From Keys to Reading. Oklahoma City: Economy Co., 1972. (a)

Harris, T. L., Creekmore, M., & Greenman, M. H. *The sun tree: Teacher's edition.* From Keys to Reading. Oklahoma City: Economy Co., 1972. (b)

Harste, J. C. Reading and reality. In J. Robinson (Ed.), *Whole language for whole children.* Columbia, Mo: Reading Department, University of Missouri, in press.

Harste, J. C., & Burke, C. L. A new hypothesis for reading teacher education research: Both the teaching and learning of reading are theoretically based. In P. D. Pearson (Ed.), *Reading: Research, theory and practice.* 26th Yearbook of the National Reading Conference. Mason Publishing Co., 1977.

Harste, J. C., & Burke, C. L. Toward a socio-psycholinguistic model of reading comprehension. *Viewpoints in Teaching and Learning,* 1978, *54*(3), 9-34.

Manual: Level 3. Glenview, Ill.: Scott, Foresman, 1971.

Menosky, D. M. *A psycholinguistic analysis of oral reading miscues generated during the reading of varying portions of texts by selected readers from grades two, four, six, and eight: A descriptive study.* Unpublished doctoral dissertation, Wayne State University, 1971.

Mitchell, K. *Patterns of teacher-student interactions during reading as related to teachers' theoretical positions*. Paper presented at the annual meeting of the National Council of Teachers of English, Kansas City, 1978.

Rhodes, L. K. *The interaction of beginning readers' strategies and texts reflecting alternate models of predictability*. Unpublished doctoral dissertation, Indiana University, 1979.

Spache, G. D. *Good reading for poor readers* (9th ed.). Champaign, Ill.: Garrard, 1974.

The three little pigs. In Scott, Foresman Reading Systems. Glenview, Ill.: Scott, Foresman, 1971.

Tolstoy, Alexei. *The great big enormous turnip*. In Scott, Foresman Reading Systems. Glenview, Ill.: Scott, Foresman, 1971.

Lynn K. Rhodes, Ed.D., is assistant professor of education, University of Colorado at Denver.

Some Basic Notions About Reading Comprehension: Implications for Teachers

Robert J. Tierney and Rand J. Spiro

Consider for a moment what is involved in reading comprehension. For example, consider what is involved in understanding the following passage:

> **Wagons Across the Nation**
>
> The journey was both exhausting and exhilarating. Some rode on horseback; others rode in wagons. The wagons wound their way across the mountains. They skirted the rivers, then they moved across the plains. We knew it would be a long time until we reached our destination.

From the title "Wagons Across the Nation," it is likely that an American reader would visualize a scenario of horsedrawn wagons traversing the United States sometime in the nineteenth century. The reader might visualize the wagons crossing mountains, then skirting rivers, and then traversing the plains. The reader might even visualize a dirty and sweaty traveler, an American pioneer, relating the tale. The direction of the journey might be assumed to be east to west, with the destination referred to being the Pacific Coast. In essence, the reader develops an interpretation that accounts for the text. Yet if the passage is reexamined, it would be noted that the text never explicitly stated that the wagons were horsedrawn, when or where the journey took place, or the components of the rich imagery that the scant descriptive details evoked. In fact, to consider just one equally valid interpretation of the story, the text could have been describing an event that was part of the recent Bicentennial celebration. In constructing an interpretation that seems plausible, it seems that a reader selects, inserts, substitutes, deletes, and connects ideas. In other words, in moving toward an interpretation that a reader deems reasonable, reading comprehension involves the simultaneous engagement of the reader's pre-existing ideas with the information presented in the text.

Two questions arise: What information is presented in the text? What information is engaged by the reader? In producing a text, an author does not merely transmit thoughts in words to the page; he also makes assumptions about knowledge the reader can be expected to share. The author may then take for granted that such information will be supplied by the reader and need not be explicit in the text. That is not to say that readers will supply only information that authors expect them to. The extent to which the reader's understanding represents the author's ideas will depend on the reader's background of experience, abilities, purposes, and reading predispositions. It may be that the reader's purposes are at variance with the author's intended purposes. For example, the "Wagons Across the Nation" passage may have been written to instill a feeling for a characteristic episode of an historic era. However a reader concerned about an upcoming test may be more interested in isolating certain "testable" facts than imaginatively experiencing the dusty nature of the trail.

The point is that the relationship between author and reader in both text production and reading comprehension is not a simple one. Instead, both text production and reading comprehension involve complex interaction among author, text, and reader; it is as if the author makes a contract with the reader and the reader makes a contract with the author. But this contract does not mean that both agree to the same terms. There are two key points, then, that should be made about reading comprehension:

1. A text is never fully explicit nor is reading comprehension exclusively textual.
2. A number of factors influence both the extent to which an author's ideas are represented explicitly in text and the extent to which a reader's understanding will vary from both the author's intended message and the explicit textual information.

Reading comprehension can then be characterized as involving and being influenced by the interaction back and forth among (a) the text, (b) the reader, including the reader's background of experience, assumptions about the author's intentions, interest, etc.; and (c) the strategies the reader employs (consciously or not) to mediate between the other two components. As Kintsch (in press) aptly wrote, "meaning is something the reader creates in response to the text, not something directly given." The text is like a blueprint that guides the *construction* of understanding (see Spiro, in press). Thus, a number of factors influence reading comprehension, and what might be considered as the accuracy with which the reader addressed the author's ideas should be regarded as relative.

Pedagogical Implications

To extend this discussion to an examination of pedagogy in the light of the postulates presented, it is suggested that if teachers understand the nature of readers' understanding, then they have the basis for structuring learning environments to stimulate and to develop a given reader's understanding, i.e., they would have a basis for determining what might facilitate and what might impede comprehension.

Along this line of reasoning, an examination of present teaching practices suggests that teachers implement a number of common practices which may be misguided. For example, teachers sometimes seem to ignore the notions that the text is never fully explicit and that reading comprehension is not exclusively textual by assuming that students should identify and regurgitate the author's ideas regardless of the text, the reader's background of experience, or the purposes for reading. Teachers sometimes disregard the notion that a number of factors influence the extent to which an author's ideas are represented explicitly and the extent to which a reader's understanding will vary, for example, by the use of strategies that elicit the same rigid interpretation of the author's ideas or require the same type of answers to the same type of questions across different texts. For example, through the use of strategies such as cloze (especially cloze with an exact replacement), teachers may require a reader to address the author's exact wording across words of varying significance and text. Through the use of strategies such as the Informal Reading Inventory, teachers often assume that the same criteria applied to the same types of questions can address a reader's understanding across different texts read for different purposes.

The key point here is that teachers often fail to realize that what is considered an appropriate understanding is likely to vary from context to context, that "accuracy of understanding" is relative. That is, criteria for "accuracy" should be considered a function of reader and text characteristics, as well as purposes of reading. It seems, then, that teachers need to understand that a "word-perfect" rendition of a text is rarely required and seems difficult to justify for any reading situation. When readers interact with text, they will and should acquire some information that was stated in the text and some information that they generate. Certainly there are situations for which it may be reasonable to expect a reader's understanding to remain close to the text; for example, when following a set of directions to build something. Alternatively, there are other situations for which it may be appropriate to expect a more personal interpretation, for example, when reading literature or poetry. To this end, teachers might address the following questions:

- To what extent was the reader's relevant background of experience activated during reading?
- To what extent was the reader's understanding adequate for the text and purposes for reading?
- When a reader's understanding diverges from some consensual author's intention, can the reader justify his or her idiosyncratic interpretation?

With these questions in mind, consider those situations when a reader's understanding may be inappropriate for either the text or purposes for reading, in particular, those situations for which readers develop an understanding which is "too text-based," "too reader-based," or "short circuited" by a lack of background experience to deal with the textual material. For example, consider the difficulties an American reader will incur when trying to understand the following passage.

Today's Cricket

> The batsmen were merciless against the bowlers. The bowlers placed their men in slips and covers. But to no avail. The batsmen hit one four after another along with an occasional six. Not once did a ball look like it would hit their stumps or be caught.

What does a teacher do in these situations? A teacher in this situation seems to have just three choices. The include: dismissing the passage as culturally irrelevant; providing the students with the appropriate experience to address the text; or providing the students an analogous framework, for example an explanation of cricket by comparing it with baseball. A key point here is to realize that the old is required to understand the new. As Rumelhart and Ortony (1977) have emphasized, "In all cases, existing knowledge is utilized in and required for the acquisition of new knowledge" (p. 117). Or as Trench (no date) commented in reference to the parables of Christ:

> He brought forth out of His treasure things new and old; by the help of the old He made intelligible the new; by the aid of the familiar He introduced that which was strange; from the known He passed more easily into the unknown. And in His own manner of teaching He has given us the secret of all effectual teaching. (p. 25)

This implies that in those text situations for which the reader lacks the background of experience, teachers need to build the reader's background or provide an analogy by which the readers can build such backgrounds for themselves.

At other times, the reader's prior knowledge is over-relied on, producing understandings that are too "loose" or "reader-based" for the text and its intended purpose. Particularly when a text deals with a familiar topic, the reader may think that he or she knows what is going to be written and therefore need not attend carefully to the text. For example, such a reader may fail to notice subtle text signals that indicate a topic is being satirized, or he or she may not detect that expectations about the text have not been confirmed. Quite often, students with an overly reader-based understanding have their difficulties compounded by the fact that "they don't know that they don't know." They presume that they know the material better than they actually do or need to. In these situations, a teacher has several options: to alert the students to the need to read the material carefully; to have the students read the material in conjunction with carrying out some activity, the successful performance of which is dependent on careful reading; to provide the students with strategies such as outlining and notetaking for carefully reading the text; to alert the students to their own level of understanding and maybe to ways to monitor that understanding; or to force the students to construct an understanding, for example, through questioning their ideas.

Finally, consider the situation when a reader's understanding is too "text based" for the text and purposes for reading. As Spiro (1977) suggests, certain conditions of schooling may predispose a reader to ascribe to text an autonomy which sponsors the separation of textual information from related knowledge which would inform its understanding. For example, a reader may minimize the interaction of his or her background of experience with a text to cope with the task of answering a series of factual questions. Here the task of the teacher is to facilitate a more integrated understanding. This facilitation might be accomplished by encouraging reader-text interaction either through discussion or through appropriate vicarious experience prior to, during, or after reading. For example, the reader might be asked to discuss his or her knowledge about a topic prior to reading a text, or would be given the opportunity to relate what is being read to his or her own experience.

The major thrust of all three of these situations, then, is that teachers should be aware that an integrated understanding requires being able to bring appropriate knowledge to bear on the text as well as to glean appropriate information from the text. It implies that there is a band of permissable understanding in terms of text and reader characteristics. Of necessity, it suggests that teachers will need to be aware of the intended purposes of both author and reader in providing strategies that facilitate an appropriate balance between reader-based and text-based processes.

In summary, some of the key points teachers should recognize are:

1. Texts are never totally explicit nor should a reader's understanding be exclusively textual.
2. Texts are not autonomous nor should it be assumed that each has a single interpretation.
3. A reader's selection, insertion, substitution, omission, and binding of ideas are not necessarily a sign of reader error.
4. "Accuracy of understanding" is misleading unless defined in terms of the author's intentions and the reader's purposes.
5. A variable balance is required between reader-based and text-based processes to cope with different texts and purposes of reading.

Finally, implicit within our discussions of the nature of reader-text interactions are two additional ideas that may have an important bearing on instructional practices and observations of readers. First, implicit within our discussion has been the suggestion that inference and interpretation are as essential to acquiring an understanding as they may be to extending understanding after reading. This idea suggests that the widely espoused notion of a continuum from literal to inferred to interpretative processes, implicit within many curriculum guides, lacks validity. Not only does it lack validity as a statement about reading comprehension, it also has questionable utility as an instructional paradigm. For example, good readers may skip literal understandings where appropriate (e.g., going directly to the nonliteral sense of "Can you tell me what time it is?"), or they may use some inferential understanding of one part of a text to facilitate the literal understanding of another part of a text. Second, implicit in our discussion has been a notion of flexibility. There are many roads to understanding and, at times, different permissable understandings. Instructional techniques that impose rigid procedures on students (e.g., study guides, structured overviews, selected questioning strategies) may interfere with approaches a reader might more naturally and effectively bring to bear, given the exigencies of text, task, and reader knowledge. Furthermore, for the same individual, the most appropriate procedures may change from situation to situation.

In summary, then, if reading teachers are to be sensitive to ways to develop a reader's understanding, they should measure the appropriateness of pedagogical principles against the desired nature of a reader's understanding. As we have suggested, doing so will require recognition of the following points:

1. A text is never fully explicit nor is reading comprehension exclusively textual.
2. A number of factors influence the extent to which an author's ideas are represented explicitly and the extent to which a reader's understanding will vary from the author's intended message and the explicit textual information.

3. A reader's selection; insertion, substitution, omissions and bindings of ideas are not necessarily a sign of reader error.

4. Texts are not autonomous nor should it be assumed that each text has a single interpretation.

5. "Accuracy of understanding" is a misleading notion unless defined in terms of the author's intentions and the reader's purposes.

6. A variable balance is required between reader-based and text-based processes to cope with different texts and purposes for reading.

7. There is not a uni-directional continuum from literal to inferred to interpretative comprehension.

8. Rigid prescriptions for how to process text disregard the fact that text, task, and reader characteristics combine to produce a variety of satisfactory approaches to comprehension in a given context.

References

Kintsch, W. Comprehension and memory of text. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*, (Vol. 5). Hillsdale, N.J.: Erlbaum, in press.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Spiro, R. J. Constructive processes in prose comprehension and recall. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: Erlbaum, in press.

Trench, R. C. *Notes on the parables of Our Lord*. New York: Redell, no date.

Robert J. Tierney, Ph.D., is associate professor of education, University of Arizona, Tucson. Rand J. Spiro, Ph.D., is at the Center for the Study of Reading, ***University of Illinois****, Champaign.- Urbana*

Bridging the Gaps in Reading: Respect and Communication

Kenneth S. Goodman

Most people in the field of reading recognize a gap between theory and practice or between research and practice. With the recent major interest of cognitive psychologists in comprehension, memory, and schema and their massive incursion into reading theory and research, another kind of gap occurs. Rather than being interdisciplinary, recent research has been multidisciplinary. Each discipline works from its own bases, uses its own methods, reports to its own constituency. There is little integration across disciplines, and gaps occur.

We could represent these kinds of gaps in a two dimensional diagram such as Figure 1.

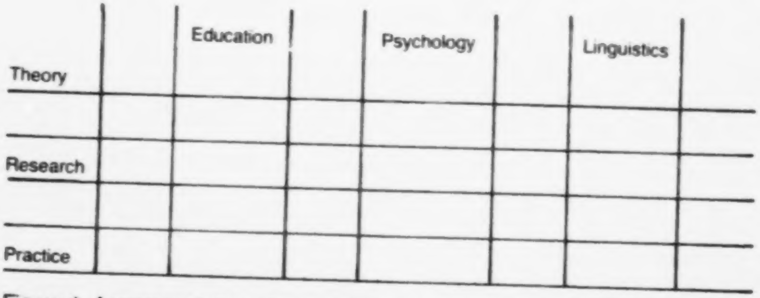


Figure 1. A representation of conceptual and methodological gaps across disciplines.

The danger of practice growing out of tradition with little input from theory and research is obvious. Perhaps not so obvious is the effect on research of having too little base in theory and practice, or theory being detached from research and classroom reality.

Similarly, it is hard to conceive of education operating without bases in psychology, philosophy, linguistics, and other social and physical sciences. But there are dangers of psychology and other foundational disciplines being uninformed by educational applications and realities.

This article will attempt to explore both the means of bridging these gaps and the problems that result from not bridging them.

The point of view that will be taken is that all dichotomies—between theory and practice, between pure and applied research, between college ivory tower and school reality, between what is ideal and what is real, between academic discipline and professional field—are artificial and non-productive. Interaction and interrelationship are essential and can improve the functioning of all. It is when "hardening of the categories" sets in and each role becomes self-perpetuating and insulated from interaction that effort becomes wasteful, duplicative, irrelevant, out of touch, and unimaginative.

It also will be argued that significant roles become neglected and lost as walls go up between professionals working in the broad field of language and reading. Ultimate purposes—the improvement of extent, effectiveness, and flexibility, of reading and other language use—get lost in a narrowing of responsibility and concern within role categories. "That's not my department" becomes the catchphrase for specialists.

At the same time specialists pressed for application of insights to teaching and learning are likely to respond with narrow direct transfer of theory and research findings to practice with no integration.

This problem of gaps developing is not unique to the field of reading or even to education. It occurs in varying forms in all fields. Consider, for example, the relationships between theoretical physics, research physics, and engineering applications of physics.

Part of the problem surely stems from specialization itself. Roles necessarily develop in every field, filled by specialists whose interactions are more likely to be with others in the same role than with those in other roles.

In reading this is certainly true. There are classroom teachers responsible for teaching reading, reading specialists of various kinds, reading researchers. But some roles have not been well-defined. It may help to consider what the continuum of roles should look like in reading (see Figure 2).

This continuum indicates five major roles. Each of these can be further subdivided into a range of roles. Though it is possible and, as has been implied, desirable for one person to engage in more than one role, they will be discussed separately.

The Role of the Theoretician in Reading

In a field like physics there is both a sharp delineation between theory and research and a continuous interaction. The modern theoretical and research physicists are different people with different work, different education, different tools. But they are in constant communication. The researcher is always operating within general (macro) theories and specific (micro) theories. The theoretician

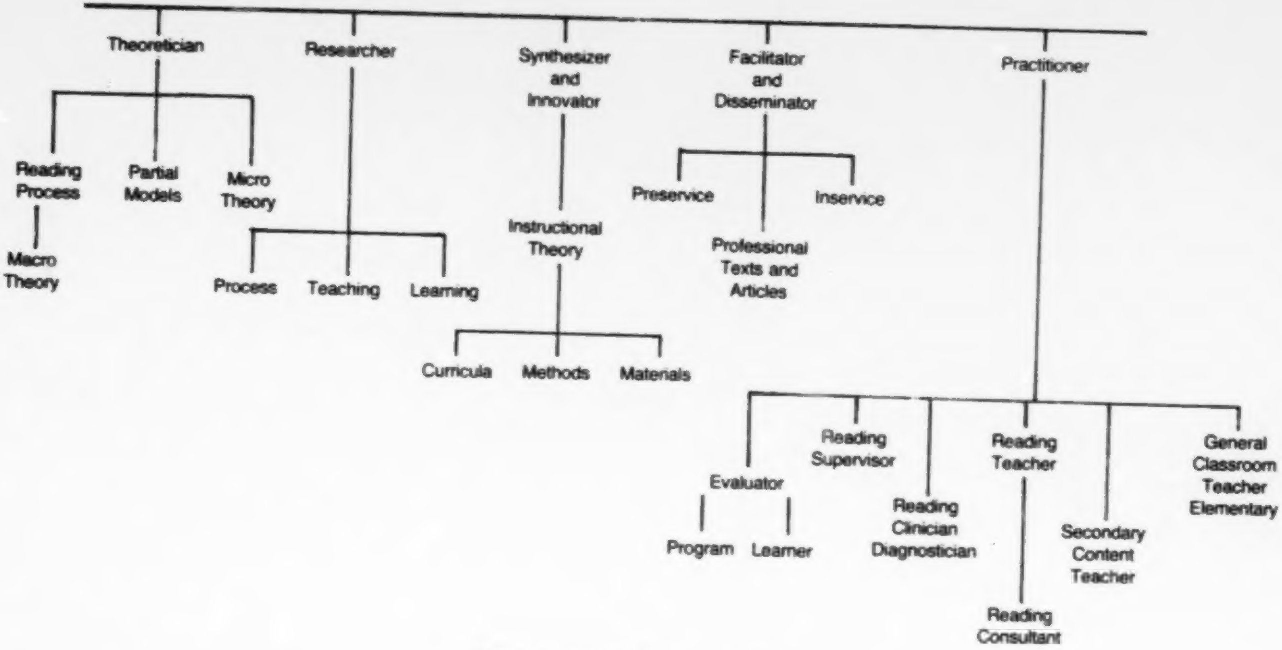

Theoretician
Researcher
Synthesizer and Innovator
Facilitator and Disseminator
Practitioner
Reading Process
Partial Models
Micro Theory
Macro Theory
Process
Teaching
Learning
Instructional Theory
Curricula
Methods
Materials
Preservice
Inservice
Professional Texts and Articles
Evaluator
Program
Learner
Reading Supervisor
Reading Clinician Diagnostician
Reading Teacher
Reading Consultant
Secondary Content Teacher
General Classroom Teacher Elementary

Figure 2. Roles in the reading field.

must constantly look to the researcher for confirming or contradictory evidence relating to his or her theory.

Often theory predicts research findings and suggests to researchers how and where to look. Often research produces unexpected findings that lead to modification overthrow or redevelopment of theory. And always practice profits from being informed about development in both research and theory. Similarly focus on practical engineering problems stimulates new directions in research and theory. The space program is a prime example.

Reading process theory was a relatively neglected field until a few years ago. It drew on theories from other fields; psychology (learning, perception, comprehension), physiology (vision, kinesics), and more recently from linguistics. Theories of reading instruction were seldom based in theories of the reading process. Partial theoretical views such as the commitment to controlled vocabulary in instruction were not theories of reading or even micro theories of aspects of reading. They generally were instructional strategies built around research findings or concepts from another field: word frequency, behavioral learning, phoneme-grapheme correspondence, to mention a few.

Goodman, in 1967, wrote a draft of a macro-theory of the reading process. It was based on research on reader's reading whole texts and on theories of language (Goodman, 1967).

Two federal efforts stimulated a great deal of interest in reading theory in the last decade. First was the USOE Targeted Research Program in Reading. The original design of this targeted research used the Goodman model as an example. The first phase of the program had as a major focus exploring the professional literature for theories and models (Gephart, 1970). Ironically most of the models and theories were generated *during* the search as researchers were stimulated to explicate their models and theoretical bases. Though the targeted research never got beyond this first phase—it died for lack of funding—it made theory respectable and brought to the forefront the need for a theoretical base for research. Some attempts such as Gough's (1972) were clearly reactive responses to other models. Still they served to publicly state and examine the theoretical base underlying some research (Geyer, 1972).

The second federal effort that stimulated work in reading theory was the National Institute of Education (NIE) program to focus funding efforts on research relating to reading comprehension. This took shape in an interdisciplinary planning conference that NIE convened, out of which came a series of panel reports. Those panel reports led to a focus on middle grade comprehension. Under the leadership of psycholinguist George Miller, a strong cognitive psychology influence emerged. Tangible results were the funding of the Center for the Study of Reading at the University of Illinois and the focus in a series of NIE requests for proposals on comprehension and on reading process.

This coincided with a broad shift in psychology from behavioral to cognitive views and a new respect for theory and modeling. Huey (1908), Bartlett (1932), and even Dewey have been rediscovered. Schema theory has been picked up after several decades of neglect. Connected coherent texts, rather than nonsense or word lists are the objects of study, and researchers find it necessary to root their design in logic (propositional analysis), in linguistics (case grammar), or in psychological micro models.

In general the reading process theorists are the researchers themselves—though not all researchers are involved in theory building.

While the field of reading process theory has been a busy one, recently two problems have emerged. One is a reluctance of theorists to commit themselves to a general theory of the reading process. That reluctance leads to a narrowing of focus in research design and a narrowing of insight in interpretation of findings. Partly this narrowing results from the preoccupation of psychologists with topics traditional to psychology: memory, cognition, comprehension rather than reading. Partly it stems from limited knowledge of language processes and linguistics on the part of cognitive psychologists doing reading research—in general they are not psycholinguists. Partly it stems from the need, real or imagined, for control in research design which causes researchers studying text comprehension to use short, artificially constructed texts which can be related to partial models of aspects of reading comprehension but cannot be completely compared to full natural language text comprehension.

The second problem is lack of activity in the construction of theories of reading instruction. Despite the importance of federal research funding and its justification in improvement of teaching and learning, few reading researchers are seriously interested in reading instruction. They may be involved in building reading process theory but they have no such motivation to build theories of reading development or reading instruction. When pressed to find applications to instruction they make use of simplistic behavioral learning models, particularly in initial reading (Anderson, Durkin, etc.) or they suggest direct use in instruction of research findings, i.e., provide direct instruction using artificial stories with competeing schemata. The issue of instructional theory will be discussed later when I deal with synthesizing.

Theories of literacy development have been curiously lacking considering the rapid progress in developmental psycholinguistics. The works of Clay, Y. Goodman, Read, and Graves are notable exceptions.

One problem comes from linguistics. A notion popular among linguists is that written language is not language in the same sense as oral language and therefore it is not learned in the same way. In fact, some argue, oral language is innate and written language must be taught. This view, unfounded in research, has inhibited extension of language development studies to literacy.

Before moving to a look at the role of the researcher, one more point about the relationship between theory and research needs to be emphasized.

Theory in any field can not be totally based on research. It must always be out ahead of research offering a coherent explanation of a field of inquiry that links the known and the unknown and generates the questions and hypotheses that lead to research. Sometimes the theory, drawing on analogy to other fields, logic, or simply brilliant insight and intuition, is far ahead of research, and must remain so until research methodology and research tools become available to explore the theoretical premises and conclusions.

Einstein's contributions to theories of the physical universe are examples. Here is his own statement on the subject.

> There is, of course, no logical way leading to the establishment of a theory, but only groping, constructive attempts controlled by careful considerations of factual knowledge. (Einstein, quoted in Eysenck & Eysenck, 1969)

In the field of reading, where much of the theory that exists comes from researchers, there has tended to be an overcautious rejection of theory not closely tied to research. This is curiously anti-intellectual and probably results from the phenomenological, logical empiricist views of behavioral psychology.

Some people would argue that linguistics, particularly that school of linguistics centered "at the sinister bend of the Charles River in Cambridge, Mass." as Henry Lee Smith once put it to this writer, has gone too far away from research in theory construction, building theory on theory with only introspection for verification. Still the productivity of linguistic theory has been immense and rapid. What is needed, as will be argued throughout this presentation, is communication and interaction between theory builders and researchers on the one hand and across disciplines on the other.

The Role of the Researcher

In discussing current relationships between research and theory, research on the reading process has already been dealt with to a considerable extent.

Research in reading can, broadly, have any of three purposes:

1. to study reading as a process,
2. to study reading development, that is, how people become readers, and
3. to study reading instruction.

Clearly there is a current renaissance of research on the reading process.

In the past, research has tended to focus on reading instruction. This, perhaps, reached its peak in the USOE sponsored "first grade studies." Experimental design was used in about 30 studies comparing different instructional programs. Popular wisdom says that these studies "proved" that differences between teachers are more important than differences in programs. While that is probably true, a more important result of these studies is what they showed of the futility of such studies. Without strong bases in competing theories of reading instruction, with little control over program differences and program implementation or even criteria for categorizing differences with only standardized tests to judge effectiveness, such studies reduce research to a "let's see what works" view. That's another term for "trial and error."

Research on how people learn to read, how they become literate, has, as indicated above, been neglected. But a good beginning has now been made in theoretically based research on the pre-school roots of literacy and aspects of acquisition (Goodman & Altwerger, in preparation; Clay, 1972). Of course, it is not easy to separate the effects of instruction from the process of learning even in looking at pre-school reading. Durkin (1966) notes that children who can read when they enter school frequently have a sister about two years older who likes to play school.

The advent of large numbers of psychologists and others entering reading research is a mixed blessing. It brings an infusion of new ideas and new talent, but it brings people in with little or no knowledge of the realities of teaching and learning reading. Two way communication between teachers and researchers is one solution. Another is that some researchers should be recruited from the ranks of experienced educators. Perhaps the best solution combines both, with research conducted by teams that include teachers, supervisors, and administrators, as well as professional researchers.

In recent years program evaluation has become confused with research. Perhaps that is because so much research relies entirely on collection of standardized test data. That leads to the mistaken view that any time test data is obtained—particularly pre- and post-tests with gain scores—then research is involved.

An example of that confusion was the follow-through evaluation. In spite of enormous complexities of program classification, controls of implementation, inadequate and incomplete data sets, etc., etc., many, including the funding agencies, are arguing that "research evidence" has shown something about effective programs for literacy and general achievement.

Methodology in reading research has tended to be narrow and unimaginative. Most of it was confined to an experimental paradigm. Issues about classroom experiments with competing instructional programs have been raised earlier in this paper. Laboratory based experiments, though more tightly controlled, have proven virtually sterile in begetting any useful insights either for theory or practice. They pulled phenomena out of written language to gain rigor and neatness of design and in so doing lost all ecological credibility.

With the shift in psychology toward concern for cognition, comprehension, and memory, and the shift toward longer (but not yet long) texts, reading research has experienced more creative research methodology with more ecological validity and more relationship to theory and practice. Research more concerned with qualitative factors (as compared with quantitative) also has begun to emerge. Ethnographic methods developed in anthropology are being applied increasingly and are getting more attention. Still, miscue research and some uses of cloze research are the only attempts to look directly at real reading by real subjects of real texts as it is taking place.

What is still needed is a wider range of reading research methodology and cross disciplinary interaction to create new forms and adapt forms borrowed from other disciplines. That goal is difficult to achieve if psychologists feel constrained to "do psychology," linguists to "do linguistics," and ethnographers to "do ethnography." To repeat one of our two refrains, we need interdisciplinary research, not multidisciplinary research.

The Role of Synthesizer and Innovator

Industry recognizes that considerable effort and expense must be devoted to integrating and synthesizing new concepts and insights from theory and research to create new applications. This is more than "product development" since often whole new industries can result. The electronics industry provides many recent examples.

In education the role of synthesizer and innovator is rare indeed. School systems rarely provide within their budgets for these roles. No doubt limited finances is a major reason. But another is that change is risky. Schools tend to be conservative or at least to wait until a pronounced trend is established before showing interest in innovation.

One major federal effort has been devoted to developing the role of synthesizer and innovator. This reference is to the establishment of educational laboratories and research and development centers. These were conceived as playing the role, which neither the school systems nor the publishing industry were playing, of translating ideas into practice, pioneering truly innovative methods and materials. The virtually total failure of this effort paralleled a tragic period in curriculum development generally in the United States. Preoccupation with technology, system, and structure dominates the efforts of these federally supported labs and centers as this preoccupation has dominated curriculum in general. Not a single one of the efforts by labs and centers in reading provided any important synthesis or imaginative use of theory and research. Not a single one had any broad acceptance or substantial impact on practice. Often instructional programs developed at labs and centers ignored research going on at the same institution.

Some federal efforts have been centered on the notion that the best use of funds would be to find exemplary practice and then package and disseminate it. There is a kind of common sense logic to the idea of finding out what good (i.e., successful) teachers are doing and then teaching others to do the same thing so they too will be successful. But implicit is the view that practice, not theory or research, should be the base of improvement of practice.

The Right to Read program several years ago selected 222 programs as exemplary and spent quite a lot of money on dissemination of 25 of these. The major criterion for selection was pre- and post-normative test results. The programs covered a wide range of current practice. No attempt was made to discover reasons for success or to question the narrow base for judging success. It is not surprising that no breakthrough resulted from this simplistic exercise in building practice on practice.

The concept of Follow-Through in its origin held more promise: synthesize competing programs on different theoretical bases and then choose the more successful ones. The process of determining success (evaluation) proved to be the destruction of this effort. It eliminated important differences, making bad programs seem effective and rendering good ones impotent.

With no examples to point toward of any successful major effort to encourage syntehsis let us turn our attention to defining the domain of synthesis and innovation.

The principal role of synthesizers in reading should be to build a sound theory of reading instruction. Such a theory must draw on sound theory of the reading process and of reading development. But it must also utilize information from learning theory, from curriculum, from child development, from all the fields of knowledge that must be synthesized in developing an understanding of what will contribute to development of literacy in the classroom.

Such synthesizers must be familiar with theory and research, with children, with schools, with current and past practice. They must build through synthesis the base in instructional theory that will make innovation in curriculum, materials, and methodology possible. The word *synthesis* has been chosen carefully here because this is not simply a matter of mixing or blending ideas and insights. The new whole that emerges must be much more than the sum of the parts.

The role of synthesizer must be performed by educationists or by gap-bridging teams centered on educationists. It requires the application, in the fullest sense of that word, of foundational concepts to reading development.

Few advanced degree programs in colleges of education or elsewhere in universities are organized to facilitate students developing the interdisciplinary breadth and depth so that they are prepared to become innovators and synthesizers. Those graduate students who acquire background in linguistics and other foundational areas tend to move into research where the prestige is and where the rewards of academia are more likely to be found.

This writer believes that the synthesizers in reading must be educationists because only people whose primary, permanent concern is education will sustain the interest in the task of synthesis long enough to create the solidly based innovations in curriculum, method, and materials.

In the absence of truly qualified synthesizers, we find diletantes from linguistics, psychology, and other fields quite willing, when tempted with royalties or consulting fees, to make ill-considered narrow applications of their own research and theory to reading. Since they know that these applications are of no interest to their disciplinary colleagues they do not exercise the same professional care in these sorties into application. The results often are naive rediscoveries of old wheels.

Alternatively some professionals within education will be impressed with new concepts from theory and research and attempt to apply them without fully understanding them or completely considering their pedagogical implications. Recent examples abound of misconceived attempts to plug in concepts from research in brain hemispheres, comprehension, memory, schema theory, cognitive style, dialects, text analysis. Often these are sincere, if superficial, efforts. But sometimes they look simply like attempts to capitalize on hot, "glamorous" ideas.

William James (1939) raised a caution about the direct utility of research findings many years ago:

> You make a great, a very great mistake, if you think that psychology, being the science of the mind's laws, is something from which you can deduce definite programmes and schemes and methods of instruction for immediate schoolroom use. Psychology is a science, and teaching is an art; and sciences never generate arts directly out of themselves. An intermediary inventive mind must make the application, by using its originality. (pp. 7-8)

True synthesis must deal with a number of difficult issues that include:

1. The relationship of teaching and learning. Too often it is assumed what is learned must be taught. Yet research on language learning suggests that language cannot be taught in any direct sense.
2. The relationship between how language works and how it develops. Over and over, we make the mistake of thinking that a good reading curriculum is based on analyzing what successful readers do, breaking that down into component skills, sequencing those, and then administering them to children. Yet nothing in research on language learning supports that view.
3. How to interpret research and theory and translate them into teaching and learning.
4. How cultural, social, ethnic, and economic differences affect acceptance of literacy—or anything else schools seek to foster.
5. How to organize instruction and package materials in order to break away from the limitations of traditional texts, lessons, and school organizations. Tradition is a terrible anchor on progress. Ultimately what is involved is shifting away from texts to the informed teacher as the responsible director of learning. That shift will involve attitude changes.
6. What the intrinsic motivations for language learning are and how to build control of language form within the context of function.

In a most real sense, if a pun can be pardoned, synthesis is where it all must come together.

For a short period when profits were high and prospects of new federal and state money seemed bright, commercial publishers appeared to be willing to make some bold moves and to encourage truly innovative programs in reading. It appeared that we were embarked on a rational path in which new text programs, if they did not lead the way, would at least follow the innovative leaders.

But economic recession, profit-motives, state and federal guidelines, and the back-to-basics movement abruptly ended that movement. Currently publishers compete in their ads to see who can sound most reactionary and archaic.

The Facilitator and Disseminator

The role of the facilitator-disseminator in reading is no less neglected than that of synthesizer. In theory every professor of education, every professional text author, every school reading supervisor or curriculum director is a disseminator. In fact, most are not. Instead, there is a strong tendency for them to be mired in the past, inhibited by the risks of advocating change, overloaded, or simply uninformed about promising innovation.

It has been traditional in many fields to couple the terms research and dissemination. The latter implies the broadcasting or distributing of knowledge and applications to potential consumers. But for this information to be useful the role must involve interacting supportively with teachers. That means facilitation, not simply dissemination.

The role this writer sees for the facilitator-disseminator cannot simply be that of purveying someone else's instructional package, bag of tricks, or recipe book. Rather, it has to be working directly with teachers to help them understand ideas, build new instructional models of their own, and then carry them into practice. The facilitator-disseminator is the crucial link between everything else and the classroom.

Such a person must have a number of attributes:

1. Interaction with theory, research, synthesizers.
2. A strong, integrated knowledge base.
3. Personal coherent models of reading, reading development, reading instruction.
4. Personal knowledge of school and classroom realities. It is not likely that such knowledge could be acquired without teaching experience.
5. Strategies and techniques for helping teachers develop insights, concepts, and applications by building on their interactions with children. This is largely a matter of gradually changing attitudes.
6. Ability to practice what he or she believes and to demonstrate the beliefs through interactions with other educators.
7. Respect for teachers and learners. This last point may be self evident but it is surprising how often those involved in working with teachers start by rejecting them. There are incompetent teachers. Most teachers, however, are dedicated, hard-working, and desirous of becoming more effective in working with children.

If teachers are made to see inservice and other dissemination activities as an attack on their professional competence, then an adversary situation is created and teachers feel a need to defend themselves by actively or passively resisting change. They develop a vested interest in maintaining the status quo.

So the facilitators must work with those teachers who recognize a need for new knowledge. They must build together; they must be sensitive to the fears of teachers, to the need for teachers to have time to think and learn and change.

Dissemination with preservice teachers is easier because they are less encumbered by traditions and established ways of doing things. But they lack the wealth of experiences that inservice teachers can relate to new concepts and new theories.

The author and his colleagues have found that the most useful device for helping teachers develop new insights is what they call protocols: vignettes drawn from the reality of real reading. Often these vignettes take the form of audio tapes of a child reading. But they may take a wide range of forms: a text for the teachers themselves to read, a videotape, a series of related episodes. The crucial thing is that they tap the sense that experienced teachers have of what is real and help them reconceptualize that reality.

In this method, the goal is not to make teachers think like researchers or theoreticians, but to help them go beyond that to a strong, comfortable control over a base for making both the moment-to-moment and long range instructional decisions teachers must make.

Fortunately, there is a dividend in dissemination. No one is more effective in disseminating new ideas and innovation than enlightened successful practitioners. They have the credibility with their peers of having used their decision-making criteria in the same settings and with the same populations as the teachers they seek to influence. So dissemination, carefully planned and carried through, can produce more disseminators.

These teacher-disseminators are the only ones in this whole continuum of roles who are in a position to bridge all the gaps in the field of reading and bring everything together. They know what they are doing and why; they can deal with the professional literature on reading theory and research, and they can help other practitioners to deal with it. They ask tough questions of theoreticians, researchers, and syntehsizers and answer tough questions of other teachers.

Professional texts can be strong tools for disseminating new concepts. They can be used by teachers who are ready for them as personal catalysts for clarifying and constructing their own instructional models. Such cases of total self-development, though they happen, are rare. Most people need more than books and ideas. They need discussion, interaction, demonstration, example, criticism, follow through. Even excellent texts are of little real use to most people except as an extension in depth of concepts already in the process of development.

Dissemination, as described above, is, of course, nothing more than good teaching, and good teaching requires patience, planning, and considerable support for the learners.

The fifth role in the reading field is that of practitioner. Once the only professional practitioners in the field of reading were primary classroom teachers. That is still true in many parts of the world. The major impetus to the differentiation of practitioner roles in reading has been negative: recognition that there are many children who are not learning to read easily and well. So we developed out-of-school clinics and in-school remedial programs with specialists responsible for finding out what was wrong with the learners and to provide theory or remediation. At the same time we devoted great attention to developmental programs that span from pre-school readiness to college level reading efficiency and study skills. And we developed supervisors for these programs and evaluators for them. Even the developmental programs have a negative justification. They give disproportionate attention, energy, time, and resources to reading as compared with other language processes and other school concerns. Furthermore, in our eagerness to avoid reading problems we isolate reading from other language processes and from its use in communication and learning.

Reading practitioner roles are geared to this negative orientation. The practitioners have an unavoidable vested interest in "the reading crisis": it is the justification for their jobs, and their expertise is often built around a technology of diagnosis and remediation of weakness.

Many school reading specialists are primarily test givers and package program technicians; they diagnose, pigeon-hole, and schedule pupils and materials.

This discussion is not intended as an attack on the reading practitioners in their varied roles. It is a call for a reassessment, however, of the way the reading profession is organized and its basis of operation.

This writer believes that the current state of reading practice is largely unfounded in articulate research-related theory of reading, reading development, reading instruction. In his opinion, such a base is absolutely necessary for real progress to be made and that its application will require major reconsideration and reformulation of the professional practitioner roles in reading.

The author's present view, which is by no means a complete one, of where the field needs to go is presented in concluding this paper.

There is need to put renewed emphasis on classroom teachers, both elementary and secondary, as the major professional reading practitioners. We need to support them with soundly-based preservice education and ongoing consultation so that they can help children develop reading *with* the other language processes in the context of functional use. That also means less reliance on special materials narrowly oriented to reading "skill" development and more teacher coordination of a wide range of materials through which children will develop and use reading. It means a greatly increased role for the teacher and a diminished one for the basal reader and other instructional packages.

The other major practitioner role to be seen is the evolution of the school and district reading specialists into facilitators working with teachers to support them in their increased roles. They probably will be labeled consultants, working with children only for the purpose of helping teachers to understand their strengths and needs. These consultants would deal not only with reading but also with all aspects of the development and effective use of language. They would understand learning to read theoretically and intuitively on the basis of their own experience.

Our society will still need adult educators with a strong understanding of literacy to work with those persons past school age who consider themselves nonreaders either because of lack of opportunity to learn or because of school failure.

Communication and interaction, as stated throughout this paper, are the keys to bridging the gaps in reading. Both depend on mutual respect across all roles as well as a sincere interest in cooperation to support each other. Weighing against this is the heavy tradition that produces an academic pecking order, disrespect for practitioners among college based researchers, mistrust of theory among teachers, etc., etc., etc.

It will take an organized and concerted effort to overcome these impediments to progress.

Funding agencies like NIE and USOE can help in this effort to overcome obstacles and in arousing more interest in the weakest links: synthesis and facilitation-dissemination.

But this kind of bridge-building is like the more literal kind. It has to be started at one or both ends of the gaps and be built across. Someone has to recognize the need and take the first steps. Some individuals who play more than one role—theoretician-researchers, synthesizer-researcher—are in key positions to bridge gaps and initiate interaction.

Professional organizations, particularly IRA and NCTE, and their journals, can provide important service in fostering interest in cross-disciplinary and cross-role interaction.

As with all problems, recognizing the need for bridging the gaps in reading is a necessary first step toward the solution. It is my hope that this discussion has contributed to that recognition.

References

Bartlett, F. C. *Remembering*. Cambridge: The University Press, 1932.

Clay, M. *Reading. The patterning of complex behavior*. Auckland: Heinemann, 1972.

Durkin, D. *Children who read early*. New York: TC Press, 1966.

Eysenck & Eysenck. *Personality structure and measurement*. London: Routledge & Kegan Paul, 1969.

Gephart, W. J. The targeted research and development program on reading: A report on the application of convergence technique. *Reading Research Quarterly*, 1970, *5*(4), 433-445.

Geyer, J. J. Comprehensive and partial models related to the reading process. *Reading Research Quarterly*, 1972, *7*(4), 541-587.

Goodman, K. S. Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, May 1967, *6*, 126-135.

Goodman, Y., & Altwerger, B. Reading: How does it begin? In C. Fisher (Ed.), *Discovering language in children*. NCTE, in preparation.

Gough, P. One second of reading. In Kavanaugh & Mattingly (Eds.), *Language by ear and eye*. Cambridge: MIT Press, 1972.

Graves, D. The second R. In *Balance the basics, let them write, a report to Ford Foundation*. New York, 1978.

Huey, E. *The psychology and pedagogy of reading*. New York: Macmillan, 1908. Reprinted, Cambridge: MIT Press, 1968.

James, W. *Talking to teachers on psychology*. New York: Holt, 1939.

Read, C. Pre-school children's knowledge of English phonology. *Harvard Educational Review*, 1971, *41*, 1-34.

Kenneth S. Goodman, Ph.D., is professor of education, University of Arizona, Tucson.

Methodological Concerns in Comprehension Research

P. David Pearson

The Context

Thomas Kuhn's (1970) notion of paradigm shifts is an appropriate philosophical space in which to begin an essay about the methodology of research in reading comprehension. Kuhn argued that in a science, progress is characterized less by the gradual accretion of knowledge and more by periodic shifts in the perspective or world view in which its scientists work and try to explain the evidence available. Progress is possible within the time frame in which a particular paradigm dominates a science. In fact, it is usually the case that ever-more-refined research within a paradigm leads to the situation requiring a shift to another paradigm. Good research serves the function of stretching the boundaries of its own paradigm to untenable positions. When they become sufficiently untenable, the paradigm bursts, making way for a new world view within the science.

Such, this author thinks, is what has happened in the linguistic and cognitive sciences during the past three decades. The behavioristic paradigm dominant in psychology, linguistics, and education generally was challenged by the data emanating from its own research. Parameters had to be extended to ludicrous ends (consider the complex Markovian chain explanations of linguistic performance prevalent in the decades of the fifties and sixties) to explain any sort of complex human activity. Hence, the nativist views of language competence espoused by Chomsky (1957), and the information processing models of several psychologists (most notably Miller) were welcomed by researchers frustrated by the narrow world view afforded them by behaviorism.

By the end of the decade of the sixties, the paradigm shift was complete, at least in spirit. Researchers felt comfortable studying entities like sentences rather than nonsense trigrams and quingrams. They incorporated variables like imaginability, grammatical complexity, and semantic cohesion into research designs. This new spirit of adventure even led some researchers in the early seventies to begin to tackle entities like paragraphs, stories, conversations and whole discourses.

This work was supported by the National Institute of Education from two sources: The work was begun at the Center for the Study of Reading (Contract No. US-NIE-C-400-76-0116) and concluded at the Wisconsin Research Development Center for Individualized Schooling (Center Contract No. OB-NIE-G-78-0217).

Journals arose that were sympathetic to these new researchers, bearing titles like *Cognitive Psychology, Cognition and Memory, Discourse Processing*, and *Journal of Psycholinguistic Research*. Established journals altered editorial policies to accommodate these new endeavors.

In the early stages of a science, or in the early stages of a new paradigm, the members of the science tend to adopt a pioneering spirit. They rationalize the lack of rigorous control of independent variables and the inclusion of mysterious dependent measures in the name of trailblazing. Researchers who use a *post hoc* empirically determined criterion for discriminating between high and low levels of contextual constraint (instead of independent a priori criterion) are not "sloppy" or inclined to circular reasoning; rather they are "breaking new ground."

As we enter the decade of the eighties, surely we will want to argue that research dealing with linguistic and cognitive behavior is emerging from its infancy into, at the least, a "toddler" stage of development. Hence, we will make greater demands on the kinds of problems studied, the types of methodological tools used in those studies, and the kinds of conclusions and speculations offered to explain the evidence collected.

This essay will raise several issues related to the methodology used in research studying the processes humans use to understand and remember written discourse. It also will deal with the relationship between research on the basic cognitive processes involved in these encounters with discourse and the research conducted to determine what educators can do to facilitate the growth of these cognitive processes. Implicit in all these arguments is the assumption that it is important that we move toward more rigorous methodology so that we may conduct the kind of research that will stretch the boundaries of our new found paradigm to the point where it will seem as ludicrous to a new generation of scientists as the old paradigm seemed to us a decade or two ago. Ironically in doing our best to validate the paradigm within which we work, we will also be doing our best to lay the foundation for its destruction.

The set of issues related to discourse comprehension research is divided into three categories that ought to be studied: issues related to the *Domain of Discourses*, issues related to the *Design and Implementation of Experiments*, and issues related to the *Role of Theory in Basic and Applied Research*. After dealing with each of the issues subsumed by these three categories, the essay will conclude with some speculations about the extant and potential contributions of research conducted within this new tradition.

Domains of Discourse

The title of this section of the paper is intended to subsume that set of issues related to the generalizability of findings from research in comprehension; that is, over what domains of real discourse can we expect the results to hold? The problem can be conveniently decomposed into two related issues: (a) the generalization to a population of linguistic elements and (b) the ecological validity of experiments.

Generalization to a Population of Linguistic Elements

Think back to the first inferential statistics course you ever took (or might have taken if you never did). Conjure up the notion of random sampling. Remember why it was important to draw a random sample of subjects from a population: So that an experimenter could feel confident that the behavior of the sample of subjects selected was typical of the behavior of the entire population of subjects. In general, researchers want to make statements about populations of subjects (such as all college students, or all adults, or all primary grade children) rather than about the particular samples they happen to have selected for particular studies.

Remember also that the statistical analysis procedures used to draw inferences from research studies take this desire into account by regarding *subjects* as a random factor in the analysis and a random effect in the underlying statistical model. Now one can question just how randomly actual samples happen to be drawn, but the research community nonetheless behaves as if it did draw random samples of subjects and uses statistical analysis tools consistent with that behavior.

Now consider the issue of linguistic elements in studies dealing with the comprehension of textually presented information. Intuitively, one would suspect that researchers are interested in drawing inferences about words or sentences or paragraphs or passages *in general* rather than the *particular* materials used in a particular study. After all, who cares if sentence 12 in a passage is harder to understand than sentence 13? What we do seem to care about is knowing that sentences like sentence 12 are harder to understand than sentences like sentence 13.

Reasoning by analogy, then, we should expect that comprehension researchers would regard linguistic variables like words, sentences, paragraphs, and passages in the same way that they regard subjects. That is, (a) they should be concerned about drawing random samples of linguistic elements from identifiable populations, (b) they should treat such elements as random factors in their statistical analysis plans, and (c) they should conceptualize them as random effects in their underlying statistical models.

Researchers, however, rarely do any of these three things. More typically, they sample linguistic elements by carefully selecting contrasting values (e.g., high- versus low-frequency words, grammatically simple versus grammatically complex sentences, attributional versus covariance passage structure). Furthermore, they often create materials rather than truly sampling from any existing population whatsoever (some have suggested that many comprehension studies may generalize to the population of passages that researchers write especially for comprehension studies). Also, researchers seldom regard a factor like passages or sentences as a random factor in the statistical analysis; if they take such a factor into account, they are likely to treat it as a fixed factor. This behavior has the net effect of conducting the statistical analysis as if the researcher was using the *entire population* of the linguistic stimuli in question.

What recourse do comprehension researchers have? How can they legimately generalize beyond the linguistic stimuli within the experiment? There are several options. Coleman (1964), Clark (1973), and Santa, Miller, and Shaw (1979) have dealt with this issue at length. The concensus among them seems to be that in situations in which linguistic stimuli do represent a sample from at least an implicit population and the researchers do want to make a statement about elements beyond those used in the experiment, then they should use a statistical model which regards linguistic stimuli as a random rather than a fixed effect. The usual suggestion (Clark, 1973, or Santa et al., 1979) is that the researchers use quasi *F* ratios rather than the simple *F* ratios derived from a standard analysis of variance (see Winer, 1971). By using the quasi *F* ratios, researchers can approximate an exact test for a situation in which their underlying model has random effects for both subjects and linguistic stimuli. The procedures for carrying out such an analysis are specified by Clark (1973), using the advice of Winer (1971).

Researchers pay a price for legitimizing their desire to generalize to a population of linguistic elements, however. Other things being equal, quasi *F* tests will tend to be more conservative than standard *F* tests, thus making it more difficult for the researchers to validate a significant effect for a given variable like, for example, sentence complexity.

There are alternatives to the quasi *F*. Researchers can be content not to generalize beyond the linguistic stimuli in the study, stating quite simply that they found such and such an effect for these particular stimuli. Intuitively, many of us are likely to regard their results as uninteresting and weak in the face of such a limitation.

Also, researchers can use the principle of replication to help them validate the strength and generalizability of an effect. Suppose Researcher 1, using his own materials, finds that concrete words are more easily remembered than are abstract words. Researcher 2, using different samples of concrete and abstract words, finds similar results. And so on for Researchers 3 through N. At some point, the research community has to develop a strong belief in the strength and validity of such a replicable effect, quite independent of how simple or sophisticated the statistics may be for a given study.

Finally researchers ought to use linguistic stimuli as an effect in their underlying statistical models and as a factor in their analyses even when they are not using quasi *F* ratios to estimate population parameters or even when they are sampling the entire population of linguistic stimuli (and can legitimately regard the linguistic factor as a fixed effect). In other words, researchers should not necessarily sum over items of a particular kind to create a summary score. To illustrate, suppose that a researcher has two treatments in which he is interested—complex sentence structure versus simple sentence structure. Then suppose that 12 bits of content (facts about the world, say) are written in each of these two treatment conditions. Two independent samples of 20 subjects are randomly assigned to treatment conditions. Finally, suppose the dependent measure is some semantic cohesiveness score (e.g., how many facts about the world are bundled together in the average sentence in free recall of sentences presented in the list). One procedure would be to give each subject a score, summed over the 12 sentences, and then to compare the mean of the two treatment groups. The statistical analysis would be a simple one-way ANOVA with a treatment factor called sentence complexity. An alternative procedure would give each subject 12 separate scores, one for each sentence. Then the analysis would be a two-way ANOVA with sentence complexity as the between-subjects treatment effect, sentence as a repeated measurement effect, and an effect due to the interaction of sentence complexity and sentence. If the interaction effect in the ANOVA proved to be non-significant, the researchers could be confident that whatever treatment effect existed generalized at least across the 12 sentences used in the study. If the interaction effect proved to be significant, then conclusions about the treatment effect will not even generalize to the sentences in the sample used in the study. In short, by including a stimuli effect in the underlying statistical model and analysis, the researchers will be able to exhibit confidence in their treatment effect in precisely the situation in which that confidence is warranted.

But suppose the researcher uses a dichotomous (right/wrong) rather than a continuous scale for scoring sentence responses. He finds himself in a bind because he is now using ANOVA procedures for 0-1 data (i.e., the scores on each sentence are either 0 or 1). Now he has to satisfy a whole set of additional assumptions in order to justify the use of ANOVA. He has a way out however. He can arbitrarily divide each set of 12 sentences into 3 groups of 4. Now he has 3 scores for each sentence type and need not worry about the additional assumptions necessary for 0-1 ANOVA. The analysis would be identical to that in which he had 12 different semantic cohesiveness scores except that in this case he would have 3 correctness scores varying from 0-4.

In summary, researchers in the comprehension domain, dealing as they do with linguistic stimuli of various sorts, are forced to deal with the issue of the domain or population of stimuli to which the results of their studies will reasonably generalize. In the past, many of us have used inappropriate if not illegitimate design and statistical procedures to draw the inferences we have. In the future, we need either to adjust our statistical models and procedures used to draw those inferences, or we need to hedge our conclusions so that they do not outweigh our methodology.

Ecological Validity in Comprehension Research

In the last half-decade it has become commonplace to denigrate much of educational and psychological research for its lack of ecological validity. After Bronfenbrenner (1976), many researchers have argued that research which fails to account for the variety of contexts (social, political, economic and linguistic) in which the education of children occurs is likely to be irrelevant in helping practical educators make decisions about educational programs.

The proverbial whipping post in such a critique is the classic psychological experiment conducted in a university laboratory using atypical populations of subjects, elaborate equipment, and dependent measures far removed from the kinds of activities children typically pursue in classrooms. However, research conducted by educators—perhaps even within a school environment—fares only slightly better. It can be criticized on similar grounds. The materials used in the experiment may not be typical of those that children really read (they may even have been rigged by the experimenter). Students may have been tested or taught individually rather than in their usual instructional groups. The task itself may have induced the children to adopt some unusual strategy unique to the experimental setting.

Criticisms of this sort put researchers into a classic dilemma. Real classroom instruction occurs in an environment in which many variables are hopelessly confounded with one another. The best teacher, using reading series A with good students from a high socioeconomic level, may use a lot of higher cognitive level discussion questions whereas the worst teacher, using reading series B with poor students from a middle socioeconomic level, may use a lot of lower level discussion questions. In such a situation it is impossible to sort out the specific factors that might account, say, for differences in achievement or attitude. Empirical research, by its very nature, is in the business of unconfounding and controlling as much variation as precisely as it can. The instant researchers move toward precision and control of variation in the environment, they create a somewhat unnatural ecology for their experiment. The instant that they move toward creating an experimental ecology as similar as possible to that operating upon a real classroom, they lose control over their basic tool: precise and controlled variation of potentially important factors.

Before discussing options available to researchers to cope with this dilemma in general, this essay will highlight some particularly gnawing aspects of the ecological validity argument as it relates to comprehension research in particular. Subsequent attention will return to the issue of solutions for both the specific problems and the broad dilemma.

Purposes and goals. Put yourself in the place of a subject in a typical experiment dealing with comprehension or semantic memory. You walk into a strange classroom or laboratory room. You meet a strange experimenter who tells you to read a passage as normally as possible (as if you were reading the morning paper) because later on you will talk about it. You read. Then you are asked to count backwards by threes from 301 for one minute. Then you are told to write down every thing you can remember from the passage as accurately as you can.

The experimenter was trying to help you establish a *purpose* for reading—a normal, everyday, "see-what's-new" purpose. But because it was an experiment (especially if you have been through one or more before), you had a different expectation perhaps in spite of what the experimenter suggested. Your *goal* may have been to read carefully, almost as if you were studying for a test.

In such a situation the experimenter will have great difficulty drawing inferences about the effect of purpose for reading on subsequent recall because of the powerful and unpredictable effect of subjects' goals in an experimental setting. Spiro (1977) has suggested just such a special strategy phenomenon. Most reading in everyday situations is done to "update one's knowledge," to integrate what is new with what the reader already knows to be true about the world. In such a situation readers are not likely to pay close attention to passage details, concentrating instead upon the state of current knowledge that exists after reading. Hence the distinctions between what was in the text and what readers knew before reading the text are likely to blur. In short, the readers usually try to integrate new information into existing schemata (or knowledge structures). In contrast, the reader who is studying for a test (or trying to remember the exact word of linguistic stimuli in an experiment) is probably better off to avoid integration with pre-existing schemata, choosing instead to construct a special short-term schema for the information presented in the text. Hence the difficulty in drawing inferences from tightly controlled experiments to normal reading activity. (Ironically, however, these experimental conditions *may* permit the experimenter to draw inferences about normal school reading because of the similarity in mind set between a subject in an experiment and a student cramming for a mid-term examination.)

Materials. To an extent, we have dealt with the issue of materials in the section on generalizing to a population of linguistic elements. That is, for materials in an experiment to stand the test of ecological validity, they must be representative of the population of materials that exist in the environmental systems (see Bronfenbrenner, 1976) to which the researcher wishes to generalize.

But an even more common and pressing argument here is the criticism that the specially rigged materials of experimental studies are not at all like what people, adults or children, ever have to read. On the face of it, the criticism is true of many materials in comprehension experiments. Several experiments (e.g., Bransford & McCarrell, 1974; Anderson, Reynolds, Schallert, & Goetz, 1977) have used purposely ambiguous passages; others (e.g., Sachs, 1967; Bransford & Franks, 1971; Pearson, 1974-1975) have used sentences with convoluted or unnatural grammatical structures. Still others (e.g., Stein & Glenn, in press) have employed stories where critical events have been purposely misordered.

Why would researchers ever want to use such awkward, unnatural, atypical materials? The answer is for purpose of precision and control; more precisely, to avoid certain confoundings. For a moment, consider the alternative. A team of researchers may be interested in the effects of complex versus simple sentence structures on comprehension. A more ecologically valid way to approach the problem would be to go out and sample large numbers of already existing complex and simple sentences, give them to subjects to read, and then contrast mean comprehension scores. But notice that the researchers have lost control over the semantic content of the sentences; sentence content is now completely confounded with complexity. Now the researchers will have difficulty attributing any comprehension differences between the two groups of sentences to complexity because it may be the case that the complex sentences contain the most difficult concepts. By contrast had the researchers rigged the materials by writing a grammatically simple and a grammatically complex version of each sentence content, they could have avoided the possible confounding.

The task. Implicit in all these arguments from ecological validity has been a similarity criterion: the greater the similarity between a phenomenon within an experiment and a phenomenon within the environment to which the researcher wishes to generalize, the greater the ecological validity of at least that aspect of the experiment. And the criterion certainly seems reasonable as a goal to which researchers should aspire. But nowhere does the issue become more complex than in the area of experimental task.

The problem was foreshadowed in an earlier glib remark about appropriateness of the set for precise memory of textually presented facts that subjects bring to an experiment should a researcher wish to draw inferences about students cramming for an examination. In other words, there may be more similarity between experimental tasks (writing free recalls or judging recognition test items) and school tasks (answering open-ended or multiple choice test items), than there is between school tasks and the comprehending behavior of people in everyday life. Then, of course, the issue becomes a philosophical excursion: What is the real environment and what is only the simulated reflection of reality?

Another aspect of the task issue is the possibility that researchers may devise an experimental task that allows subjects to create a mind set or adopt a special strategy that is unique to the experiment and hence uncharacteristic of the behavior under study. We have already dealt with the mind set issue in discussing purposes set by the experimenter versus goals adopted by the subject. Another example: Suppose a group of researchers wishes to study children's comprehension of time relations. The task is to read two event statements set in the context of a larger story, perform some interpolated task, and then tell the experimenter which of the two events occurred first. Then the task is repeated for several more examples. It is possible that even a young school-aged child may catch on fairly soon. He might decide to make an educated guess about the two critical events (they might even be recognizable by some textual feature, such as the sudden inclusion of time cues like before, after, then, or earlier), ignore the information in the rest of the text and the interpolated task and concentrate his mental efforts on rehearsing the order of the two events. Thus his score on the time relations task would be an inflated index of his typical ability to remember such items in a story reading situation.

Some solutions. This essay has only scratched the surface of the ecological validity issue. In fact it has dealt with issues that Bronfenbrenner (1976) would likely place within that aspect of the environment that he calls the micro-system. The larger environmental contexts such as the meso-system, the exo-system, and the macro-system that Bronfenbrenner considers so influential and essential for researchers to study have been avoided. By way of justification, only modesty of intent, interest and expertise, can be cited.

Of three specific issues raised—purposes versus goals, materials and task—the first and last are easier to resolve than the second. Related to purposes and tasks, I suggest that the similarity criterion should guide our experimental ingenuity. We should try to use tasks and mind sets as similar as possible to those that we believe exist in the environment about which we wish to draw inferences. Bronfenbrenner (1976) has made a suggestion that should help researchers, especially with the issue of subjects' goals in an experimental setting. In discussing phenomenological validity, he has demanded that researchers assess each subject's "definition of the situation, how he or she perceives the setting and its various elements" (p. 8). In a positive vein, more and more researchers are asking and reporting what subjects believe was happening in experiments.

With regard to materials, this author thinks that critics of empirical investigations have overstated their case. Quite often researchers have to rig materials simply to evaluate the power of a variable. The alternative—finding existing materials that exemplify the values of the variable in question—is fraught with such imprecision and such high probabilities of variable confounding that it presents great threats to the internal validity of comprehension experiments. Besides, this author has never been able to understand what is so sacred about "existing materials that people really have to read." Presumably they were originally written by another mortal.

But these specific problems ultimately involve the larger dilemma set forth at the beginning of this section. The best traditions of the scientific method push researchers toward precision and control of potentially relevant variables. Those criteria push researchers toward the laboratory environments, a setting in which they can vary one factor at a time while all others are held constant. The goal of making a difference to society, on the other hand, pushes researchers toward research into more real-life settings in which precision and control are made difficult.

Some resolution of the dilemma is possible if we adopt a stage-wise model of research, similar to that suggested by Hilgard and Bower (1975). They suggest a continuum of stages from basic research about behavioral processes at the one end to advocacy and adoption of technological advances at the other. A researcher who wishes to understand how skilled readers can understand and remember what they read is justified in using measures such as reaction times to various recognition probes to determine the saliency of various kinds of text statements. But such a task may be irrelevant or even misleading to the curriculum designer who wishes to design comprehension questions that maximize children's understanding of social studies material. At a fairly basic stage of research, when a researcher wishes to know whether a particular variable is operative at all, rigged materials probably will provide the most efficient path toward making that decision. At a more applied stage, researchers will surely want to confirm the hypothesis generated from such an experiment using materials "that people really have to read." Indeed Bronfenbrenner recognized the practical value of basic research with just such a continuum implicit in his reasoning (1976, p. 9).

Laboratory studies have the advantage of being more readily amenable to rigorous experimental control. Their reduced ecological validity, however, seriously limits the generalizability of results to real-life settings. Accordingly, in educational research, their value is for the exploration and clarification of hypotheses, *not* for definitive testing.

Design and Analysis of Experiments

As is the case with most domains of study, separating issues by topic is but an arbitrary convenience for both author and reader. As we move from domains of discourse to design and analysis, this arbitrariness is thoroughly exposed. In point of fact, design and analysis have been discussed here all along. However the author has chosen to highlight three personally important issues: replication, multiple dependent measures, and design complexity.

Replication

It is perhaps an inevitable part of the research tradition that the conclusions from a study remain a part of the conventional wisdom long after the details and qualifications are forgotten. Further the author's experience suggests that we are too accepting of our colleagues' findings.

Often when we fail to understand exactly what was done or what was found or what data supports a particular conclusion, we attribute the blame to our own lack of competence (or time, if we are in a hurry) rather that to the author of the article. Somehow data get lost, but the conclusions live on, regenerating themselves.

To cite an example: It is rare to pick up a paper dealing with the influence of context on word recognition without finding a reference to Goodman's (1965) article, "Linguistic Cues and Miscues in Reading." It has become a classic. For about four years, Minnesota graduate students and I have been trying to replicate Goodman's data on which is based the popular conclusion that students can read correctly in context some 60 to 80% of the words they cannot read when the same words are presented in isolated lists. We have not been able to replicate Goodman's findings.

We have consistently found that our data lag some 20 to 30% behind Goodman's. For example, he found that first grade students could correctly read in context 60% of the words misread on isolated lists (for third grade the comparable data was 80%). In two replications, our best estimates are 40% for first grade students and 55 to 60% for third grade students. Furthermore we have found that we can get about 20% improvement just by letting students read the list twice. Why we are not sure; but we suspect it has to do with internal feedback (the student narrows the word down to a couple of choices, tries one, then the other). If our data, based on two replications, are accurate, then we ought to conclude, yes, context helps, but not nearly to the degree that Goodman's (1965) data would lead us to believe. Correcting for the improvement from sheer exposure to the list a second time, it would seem reasonable to assign a 20 to 30% benefit to context (not 60 to 80%).

Negin (1977) imposed some other restrictions on the context effect. It seems to operate at the 50 to 60% level for material about at students' instructional level; however, when students read material at frustration level, the added benefit of context drops off to 15 to 30%. Furthermore, at frustration level, students often miss words in context that they got right when they initially read a list (Goodman found that this never occurred).

Another example is from a study that the author did (Pearson, 1974-1975). While studying the effect of grammatical complexity on children's recall of causal relations, Pearson found that grammatically complex statements exhibited greater surface structure stability from input to recall than did grammatically simple statements. Furthermore, grammatically simple inputs tended to be recalled in grammatically more complex utterances. From this, Pearson concluded that grammatically complex surface forms exhibited more stability because of greater semantic cohesion. In fact, Pearson found no instances in which there was a drift from syntactically complex input to syntactically simpler recall. A year or two later, he replicated the study intact. While he found the same general trend, the effect was not nearly as dramatic as it was the first time.

The point to be made is that replication ought to be encouraged; however our present system of refereed publication systematically discourages it ("this article adds nothing new to the literature"). In fact, researchers should probably replicate their own findings at least once before they publish an article reporting the findings. Perhaps if Goodman had done so, he might have discovered the problem we found at Minnesota; alternatively, he may have strengthened his original findings. If the author had done so, perhaps he would have qualified his conclusions.

Multiple Dependent Measures

It is not uncommon to read a study whose raison d'etre is that it contradicts the findings of a previous experiment. Usually careful reading of both papers will reveal some important discrepancy; quite often the discrepancy will be in the dependent measures used to assess comprehension. One researcher used a cloze test; the second, open-ended questions.

A concrete example: Neilsen (1977) completed a dissertation in which he examined comprehension as a function of story structure (a la Thorndyke's, 1977, story grammar) and conceptual abstractness (concrete versus abstract referents for words in stories). But he decided to do three modified replications of the experiment; the replications differed in terms of the type of probe used to assess comprehension (questions versus free recall versus a Bransford and Franks, 1971, type recognition task). In all other aspects, the replications were identical. The discrepancy among the replications is remarkable. In the question probe experiment, everything had an effect; in the recognition experiment there were very few effects; the recall experiment yielded results more like those of the question probe study. Neilsen's conclusions about the effects of these variables on comprehension are appropriately qualified only because he bothered to assess the generalizability of findings across dependent measures. Had he done only the question probe or only the recognition experiment, he would have accordingly over- or understated the power of the independent variables in his study.

Design complexity

If we take seriously Bronfenbrenner's (1976) advice about the ecological validity of educational research, we will end up using more complex research designs than we do presently. We will do so because of our concern for the generalizability of our research findings.

An example: Suppose we want to study the effect of various types of vocabulary pre-reading instruction on story-specific as well as more global (a transfer task) comprehension. Our main concerns are threefold: (a) Do students understand

168

stories better when they get vocabulary instruction prior to reading? (b) Does the type of instruction make a difference for specific story comprehension? (c) Does such instruction transfer to performance on other comprehension tasks when the intervention is no longer present (e.g., new stories or standardized tests)? A simple approach to the research would be to pick highly similar groups of students (age, SES, IQ, etc. would be the same across treatment groups) and carry out the study with materials that were also similar (literary genre, readability, and content might be fixed at given levels).

However, concerns about ecological validity and generalizability would push us in other directions. Instead of zeroing in on given levels of age, IQ, SES, literary genre, readability, and content, we would head in one of two directions. We might randomly sample across potential levels of these variables, or we might include two or more levels of each of the variables in our design. In the latter case we might use 9-, 12-, and 15-year olds; high and low IQ; high, middle, and low socioeconomic levels; narrative and expository genres; at and above grade level readability; and social studies versus science materials. But now our nice simple experiment is inordinately complex. The earlier experiment had perhaps six experimental cells; the revised experiment has, at least potentially, 1,152 cells. Obviously there is some compromise between the two extremes. Somewhere there is an optimal trade-off between our concern for experimental precision (which argues for simplicity) and our concern for generalizability (which argues for complexity of design).

Reading Theory and its Applications

Basic and Applied Research

In reality there exists some mysterious and ill-understood entity called the reading process. In principle at least there exists the theory of the reading process that explains what that process is and how it works—precisely, completely, and parsimoniously. In actuality, there exist many theories of the reading process that try to explain it with as much precision, completeness and parsimony as possible.

How does a science move from the theories that its scientists build to the theory, to true understanding of the process? Presumably, research forms the bridge. Research is an evaluative device. It allows us to judge which of the pretending theories deserves to be accepted and believed based upon how well each stands the tests of precision, completeness, and parsimony. The interplay between research and theory is reciprocal. Sometimes the evaluation process is conceptually-driven; that is, a theory or an element of a theory help to determine the methodology for the research and the type of data to be gathered. At other times the evaluation process is data driven; a particular finding from a study suggests a new element for a theory, a modification for an element in a theory or an entirely new theory. Whatever the direction of the process, the goal is always the source: To move closer to the theory, a true understanding of the process itself.

When this evaluation procedure is applied to the study of cognitive process, we call it basic research. These relationships are depicted graphically in Figure 1.

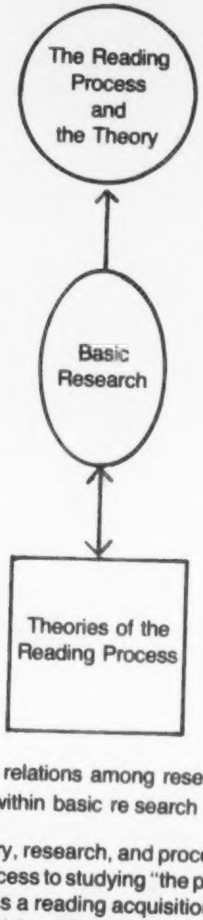

The Reading Process and the Theory
Basic Research
Theories of the Reading Process

Figure 1. The relations among research, theory, and process within basic re search in reading.

Similar relations among theory, research, and process must hold when we move from studying the reading process to studying "the process of becoming skilled at the reading process." There is a reading acquisition process, there are theories of reading skill acquisition, and there is a body of research devoted to understanding how people acquire reading skill. Figure 2 depicts such a set of relations. Presumably the same reciprocal relationship between theory and research holds here.

Surely there will be points at which there is articulation or continuity between these two domains of research/theory relations. We would expect that even a novice's reading strategies would bear some similarity to those of a skilled reader. We would expect theories of acquisition to include many of the same variables as theories of the process: purpose, text difficulty, prior knowledge, and a host of indices of individual differences.

And there will be points of discontinuity. Many variables that appear in a theory of acquisition will not appear in a theory of the reading process: distractability, time

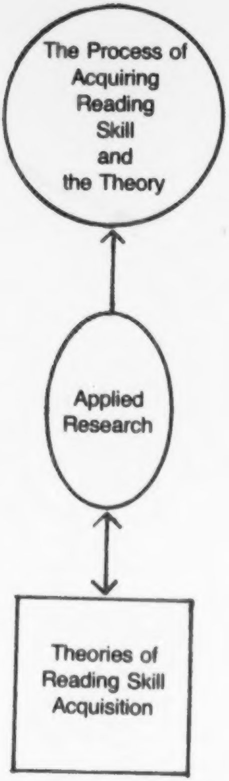

The Process of Acquiring Reading Skill and the Theory
Applied Research
Theories of Reading Skill Acquisition

Figure 2. The relations among research, theory, and process within applied research.

on task, learning rate, teacher quality, instructional method, and a host of instructional variables peculiar to the ecology of the educational setting.

If we place these two domains of theory/research/process relations side by side, as in Figure 3, we can emphasize points of articulation. In so doing, we can begin to talk in a more refined way about bridging the gap between theory and practice.

The first point that becomes obvious is that there are many "gaps" to bridge: theories of reading to the process of reading, theories of acquisition to the process of acquisition, the process to the acquisition process, basic research to applied research, theories of reading to theories of acquisition, etc.

Second, bridging is a two-way street; it is just as likely that knowledge about acquiring reading skill will inform our knowledge about the reading process as vice-versa. Too often we have assumed that we must settle issues of basic research before we can tackle issues of applied research. The model in Figure 3

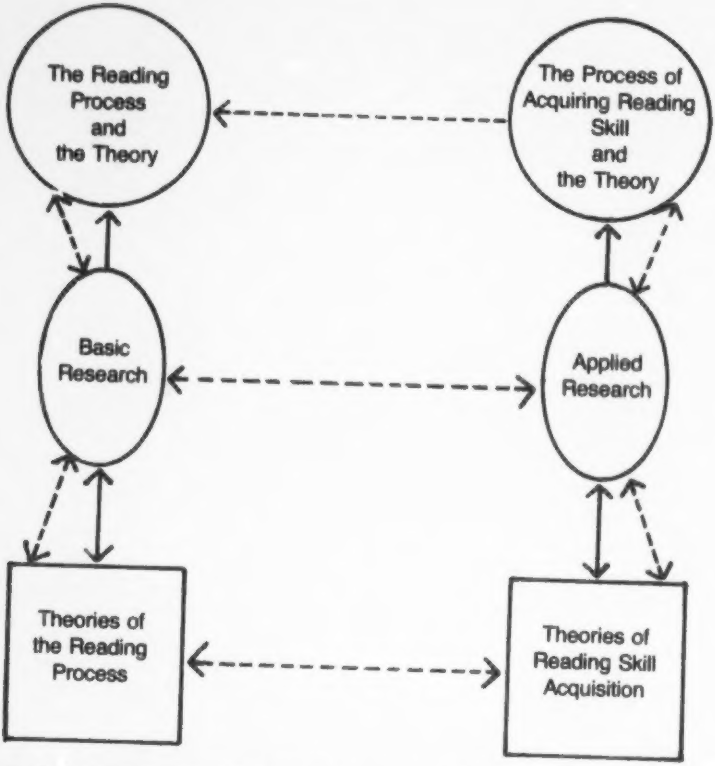

The Reading Process and the Theory
The Process of Acquiring Reading Skill and the Theory
Basic Research
Applied Research
Theories of the Reading Process
Theories of Reading Skill Acquisition

Figure 3. A specification of "bridging" possibilities in basic and applied reading research. Dotted lines indicate elements between which bridging (i.e., the one element "informs" the other) can and should occur. Note that research is a bridging medium between competing theories and corresponding processes.

suggests that such a delay in facing applied research questions may be inadvisable as well as unnecessary.

The metaphor of a two-way street carries practical implications for the conduct of research and communication requirements among professionals. Basic researchers are not, in the strictest sense, required to emphasize the implications of their research for practice (i.e., the process of acquiring reading skill). Yet, it would behoove them to speculate about "possible" implications, for then applied researchers could evaluate those speculations. Just as surely, applied researchers should speculate about the implications of their research for theories of the reading process as well as theories of the acquisition process. William James admonished society "never to question the origin of good." Surely this advice holds for the entire domain of reading research. All researchers should revel in serendipity rather than decry it.

Lest we dwell too long on the positive transfer across basic and applied domains, we need to emphasize the possibilities associated with discontinuity rather than articulation.

First, it may be a fact of life that the salient variables in the acquisition process relate to factors only incidentally related to the reading process. That is, it could be true that if we were equipped with the theory of the reading process, we could achieve a 10% improvement in our ability to improve the teaching and learning of reading. Whereas, if we knew more about the influential contexts of the home environment and/or the ecology of the school, we might be able to achieve a 40 or 50% improvement in our delivery system.

Second, it may be the case that there is a complete discontinuity between the process associated with skilled readers and the theory of the process that we ought to assume to be true as we develop instructional procedures. Let me illustrate. It could be true that top-down theories of the reading process (i.e., theories which assume that one starts with a search for meaning and uses graphic data to verify hypotheses about meaning) better describe, explain and predict what mature readers actually do. However, it could also be true that the best way to get children to read in a top-down manner is to teach them as if reading were based upon a bottom-up theory (i.e., a theory that assumes that a reader begins with decoding of print to speech and uses a speech code to mediate a link to meaning). Just as surely, the reverse could be true.

The point of this entire section is that researchers concerned with issues of practice need not and should not await the delivery of the investigation of the reading process before they undertake theory of the acquisition process. If wholesale discontinuity truly exists, such delays will set the field back even further. Even if continuity exists, practical researchers should plunge ahead on the grounds that applied research may be as useful to basic research as basic is to applied.

The Role of Theory in Educational Decision Making

It is not uncommon to hear classroom teachers decry theory. As a teacher educator, the author rarely begins a course on reading instruction without someone asking the question, Will your course be practical or is it all theory? The obvious assumption underlying such a question is that theory and practice are mutually exclusive domains, that the two do not inform one another.

It is possible that this compartmentalization of theory and practice on the part of teachers reflects little save the naivete and inadequacy of our present theories. After all, if our theories were good, they would be perceived as practical because they would describe the acquisition process and predict accurately what teachers ought to do in a host of different circumstances.

Alternatively, it is possible that the perceived compartmentalization results from our inability to communicate effectively to practitioners. Many teachers seem to have the following operational definition of theory: anything that is obtuse, irrelevant, and hard to understand. If this is true it must be a generalization derived from experience with teacher educators. We must not have communicated our theories well. Had we done so, most teachers would abide by the homily, nothing is as practical as a good theory.

That teachers do not accept that homily is an indictment of the philosophical expertise of teacher educators. We have been unable to convince teachers of two important facts about the world: (a) If teachers based their decisions on an explicit theory of reading skill acquisition, they would make instructional decisions that, at the least, were consistent with one another; (b) Even if they are unaware of it, teachers carry with them into the classroom an implicit theory of reading skill acquisition and, probably, an implicit theory of the reading process. Harste and Burke (1977) completed a clever study to emphasize this latter point. They found that they were able to categorize teachers' implicit theories of reading by observing their interaction with children in reading classes.

I do not believe that teachers have any natural predisposition to despise theory. It is a learned behavior—learned, unfortunately, at the hands of those who most want to communicate its value. What teachers need to be convinced of is the fact that theory can and does play a role in making educational decisions. For an attempt to carry out such a persuasion see Pearson and Kamil (1978) or Pearson and Johnson (1978).

References

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. Frameworks for comprehending discourse. *American Educational Research Journal*, 1977, *4*, 367-381.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331-350.

Bransford, J. D., & McCarrell, N. S. A sketch of a cognitive approach to comprehension. In W. Weimer & D. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, N.J.: Erlbaum, 1974.

Bronfenbrenner, U. The experimental ecology of education. *Educational Researcher*, 1976, *5*, 5-15.

Chomsky, N. *Syntactic structures*. The Hague: Mouton, 1957.

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 335-359.

Coleman, E. B. Generalizing to a language population. *Psychological Reports*, 1964, *14*, 219-226.

Goodman, K. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, *42*, 639-643.

Harste, J. C., & Burke, C. L. A new hypothesis for reading teacher research: Both the *teaching* and *learning* of reading are theoretically based. In P. D. Pearson (Ed.), *Reading: Theory, research, and practice* (Twenty-sixth yearbook of the National Reading Conference). Clemson, S.C.: National Reading Conference, 1977.

Hilgard, E. R., & Bower, G. H. *Theories of learning* (4th ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Kuhn, T. S. *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago, 1970.

Negin, G. *The effects of reading ability and difficulty level on students' abilities to recognize words in isolation and in context.* Unpublished doctoral dissertation, University of Minnesota, 1977.

Neilsen, A. *The role of macrostructures and relational markers in comprehending familiar and unfamiliar written discourse.* Unpublished doctoral dissertation, University of Minnesota, 1977.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall and conception of certain semantic relations. *Reading Research Quarterly,* 1974-75, *10,* 155-192.

Pearson, P. D., & Johnson, D. D. *Teaching reading comprehension.* New York: Holt, Rinehart, & Winston, 1978.

Pearson, P. D., & Kamil, M. L. *Basic processes and instructional practices in teaching reading* (Reading Education Report No. 7). Urbana, Ill.: Center for the Study of Reading, University of Illinois, December 1978.

Sachs, J. D. S. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics,* 1967, *2,* 437-442.

Santa, J. L., Miller, J. J., & Shaw, M. L. Using Quasi *F* to prevent alpha inflation due to stimulus variation. *Psychological Bulletin,* 1979, *86*(1), 37-46.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Erlbaum, 1977.

Stein, N. L., & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *Discourse processing: New directions.* Norwood, N.J.: Ablex, in press.

Thorndyke, P. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology,* 1977, *9,* 77-110.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1971.

P. David Pearson, Ph.D., is professor of education, University of Minnesota, Minneapolis.